University of Alberta

Investigating the interactions between *P* elements and *vestigial* alleles in *Drosophila melanogaster*

by

Kyle Jarvis Anderson

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Molecular Biology and Genetics

Department of Biological Sciences

Edmonton, Alberta
Spring 2008

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-45394-0
Our file Notre référence
ISBN: 978-0-494-45394-0

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

P element mutagenesis allowed the initial cloning of the *vestigial* (*vg*) allele in *Drosophila melanogaster*. While the alleles served their initial purpose, some showed an interesting phenomenon of repressor sensitivity which warranted further study. The most striking allele, $vg^{21\text{-}3}$, exhibited a near-wingless phenotype that could be completely alleviated by introducing a *P* element expressing the repressor protein. A new series of *vestigial* alleles has been created for the purpose of studying the interaction between *P* elements and *vestigial*. One allele, vg^{21r36}, is molecularly very similar to the initial *P* element allele, vg^{21}, but is much more deleterious to wing function over a null allele. Both elements are inserted at the same genomic position and have a large internal deletion. The difference in phenotypic severity allowed a focus on the small sequence differences to find that the P[21r36] element encodes a full DNA-binding THAP domain, while P[21] encodes a truncated version. The P[21r36] element was subsequently used in a test for somatic *P* element repression to show it encodes the smallest *P* element repressor to date. A second allele, vg^{21r7}, proved to be identical to $vg^{21\text{-}3}$ but with three additional P[21-3] elements inserted on Chromosome II. Tests for *P* element repression in this stock, and derived recombinant stocks, have shown that the P[21-3] element is also a potent repressor-making element. Taken together, these two repressor-making *P* elements show a much greater range of internally deleted elements can encode repressor than previously believed. Another set of experiments was designed to study how *P* repressor modifies the $vg^{21\text{-}3}$ phenotype. By monitoring levels of *P* element transcription we have shown that *P* repressor is able to downregulate transcription from the *P* element promoter. That the presence of *P* element RNA was deleterious to wing development

suggested a role of RNA interference in determining the impairment of *vestigial* in this allele. RNA folding studies determined secondary structures in the P[21-3] transcript which could potentially encode microRNAs. The $vg^{21\text{-}3}$ stock was then combined with RNAi mutants to determine their effect on wing phenotype severity. The restoration of wing development in these studies implicates microRNA production in determining the level of *vestigial* impairment in some, but not all, *P* element alleles of *vestigial*.

For my family, especially those who could not be here to see me graduate.
Your support has truly been a gift.

Acknowledgements

What a road to get to this point. Who would have guessed that someone accepted into an undergrad degree in honors physics would end up here. Thanks to my brother for pointing out that I could register for whatever classes I wanted, so I took biology instead. It was that Genetics 197 class in my first semester (taught by none other than Ross Hodgetts) that sparked my interest and set my academic path. Undergrad was eventful, I met my future wife, my Mother died of cancer, and I grew a lot as a person. Attempt to get into grad school #1 was a rejection, so I took another semester. Attempt to get into grad school #2 the next year was also rejection, so I got a job in a plant genetics lab. 2 months into that job I got a call: they needed a TA with a genetics background and were offering me a spot. My wife Erron, always supportive of my aspirations, reminded me that it didn't matter how I got in, just that I got to do what I wanted to do. Mike Russell's amazing developmental course had given me the desire to go into a *Drosophila* lab, and I was given the choice of going with Ross or John, having never really spoken to either. Thank you both for offering me a spot so I could fill the 207 roster. I came into Ross's lab and got pretty much nothing done in the first year, chasing down the mythical "intron 2 transcript". If anything, I learned how to isolate good RNA. For someone with a severe public speaking phobia, being a TA was the best thing that could have ever happened to me. Thanks to Kathy for being my 207 compatriot, every 207 TA needs someone to share the misery with. I branched out as a TA to higher level classes and came to love teaching, another thing I never would have predicted. Grad school was good, friends in the lab like Monica and Ankush made doing science fun as they were always ready to look at the latest result, listen to a new theory and help plan the next step.

Things were getting on track, I was set to be out in less than 6 years and decided starting a family didn't need to wait until after graduation. Katie, how I love you, I hope nobody reading this ever has to feel the despair of holding their dead child in their arms. Almost a year went by with barely any work done, I seriously thought about just quitting school, but once again Erron was my biggest support. In the next year a lot of little findings started to come together to help unify this thesis, things were back on track… and then the twins came. Elijah and Petra, joy of my life, thank you for fixing my heart. Finally, the last couple years of poor-to-no funding caught up with me and I needed to work as a sessional instructor to pay the bills. So here we are, i's dotted and t's crossed, I am ready to finish this degree and move on to the next chapter in my life

Table of contents

Chapter	Page
I. Introduction	1
Literature Cited	24
II. Materials and Methods	28
Literature Cited	38
III. Primary and secondary events which revert or repress the phenotype of a *P* element induced *vestigial* allele in *Drosophila melanogaster*.	39
Literature Cited	64
IV. Non-canonical *P* element repressor-making elements in *Drosophila melanogaster*.	66
Literature Cited	99
V. The suppressible vg^{21-3} phenotype is caused by a microRNA produced by *P* element transcription.	101
Literature Cited	130
VI. Conclusions	133
Literature Cited	144

List of Tables

Table	Description	Page
II-1	*Drosophila* stocks used in this study	36
II-2	Synthesized oligonucleotides used in this study	37

List of Figures

Figure	Description	Page
I-1	Sequence elements within the *P* element transposon.	18
I-2	Sequence elements within the *vg* gene.	20
I-3	Wing phenotype severity scale.	22
II-1	*P* element and *vg* PCR primer locations.	34
III-1	Phenotypic response of vg^{21} and $vg^{21\text{-}3}$.	52
III-2	Sequence of initial *P* elements.	54
III-3	Alignment of six new alleles with P[21].	56
III-4	New alleles in heterozygous combination with a null allele.	58
III-5	Southern hybridization of novel stocks created in the P element replacement experiment.	60
III-6	Mechanism of targeting P[21] into $vg^{21\text{-}3}$ to create vg^{21r4}.	62
IV-1	Inverse PCR screening of vg^{21r7} *P* elements.	81
IV-2	Alignment of recombinant markers and vg^{21r7} *P* elements on chromosome 2.	83
IV-3	PCR primers used to study new insertions.	85
IV-4	Southern analysis of vg^{21r7} recombinant stocks.	87
IV-5	Phenotype of vg^{21r7} recombinants over vg^{B} at 25°C.	89
IV-6	Expression of P[*lacZ* w^{+}] markers in combination with vg^{21r7}.	91
IV-7	P[*lacZ* w^{+}]*ci* expression in stocks carrying *P* repressor.	93
IV-8	Coding capacity of new repressors.	95
IV-9	Repression test of an insulated P[21-3].	97

V-1 Nuclease protection assay to detect *vg* and *P* initiated transcripts. 116

V-2 Nuclease protection assay to compare *P* initiated transcription with and without repressor. 118

V-3 *P* element transcript predicted secondary structure. 120

V-4 Wing area for *vg* mutants in combination with RISC mutants. 122

V-5 Representative wings from *vg* mutants in combination with RISC mutants. 124

V-6 Wing phenotype of *vg* alleles over deletions at 25°C. 126

V-7 Model of P[21-3] miRNA production and action. 128

VI-1 *P* element transcript folding from P[21] and P[21r36] 142

List of symbols and abbreviations

aa	amino acids
Ago1	*Argonaute 1* gene
Ago2	*Argonaute 2* gene
AP	anteroposterior
armi	*armitage* gene
b	*black* gene
Bc	*Black cell* gene
BLAST	Basic Local Alignment Search Tool
bp	base pairs
c	*curved* gene
ci	*cubitus interruptus* gene
cn	*cinnabar* gene
cpm	counts per minute
CyO	Curly of Oster chromosome
dCTP	2'-deoxycytidine 5'-triphosphate
Δ2-3	*P* element with 2-3 intron deleted
DNA	deoxyribonucleic acid
dp	*dumpy* gene
DV	dorsoventral
e	*ebony* gene
EDTA	ethylenediamine tetraacetic acid
g	grams

GFP	green fluorescent protein
hrp48	Heterogeneous nuclear ribonucleoprotein at 27C
kbp	kilobase pairs
kD	kiloDaltons
lacZ	β-galactosidase
LZ / LZIP	leucine zipper
M	molar
mg	milligrams
miRNA	micro RNA
mL	milliliters
mM	millimolar
mRNA	messenger RNA
NPA	nuclease protection assay
°C	degrees Celsius
PBS	phosphate buffered saline
PBT	phosphate buffered saline with Triton-X100
PCR	polymerase chain reaction
piRNA	Piwi-interacting RNA
pr	*purple* gene
PSI	*P* element somatic inhibitor
px	*plexus* gene
RISC	RNA-induced silencing complex
RNA	ribonucleic acid

Roi	*amos*$^{Roi\text{-}1}$: allele of *absent MD neurons and olfactory sensilla* gene
RpL32	*Ribosomal protein L32* gene
ry	*rosy* gene
Sb	*Stubble* gene
Ser	*Serrate* gene
siRNA	short interfering RNA
sn	*singed* gene
sp	*speck* gene
Sp	*wg*$^{Sp\text{-}1}$ = Sternopleural allele of *wingless* gene
spn-E	*spindle E* gene
Tb	Tubby chromosome
TBE	Tris Borate EDTA
TE	Tris EDTA
TESS	Tris EDTA spermine spermidine
THAP	THanatos (death)-Associated Protein
U	units
Ubx	*Ultrabithorax* gene
UTR	untranslated region
vg	*vestigial* gene
VG	vestigial protein
vgB	Df(2R)vg^{B} deletion chromosome
w	*white* gene
y	*yellow* gene

Chapter I: Introduction

The goal of my research has been to gain a better understanding of the complex interactions between a transposable element, the *P* element, and a gene which it has inserted into, *vestigial* in *Drosophila melanogaster*. Using known alleles as a starting point and creating additional *vestigial* alleles has allowed the common and unique sequences in each allele to be studied in relation to the severity of wing malformation they produce. The suppressible nature of some alleles has created interest in studying how mutant severity is determined, what is required for suppression, and how that suppression is effected. By developing an understanding of how this transposon and gene interact we may better understand how a transposon transitions from a genomic invader to a benign, and possibly even beneficial, addition to the genome.

Transposable elements

In every animal, plant and bacterial genome studied, transposable elements can be found - mobile DNA that is a component of the long-termed "selfish" or "junk" DNA from its lack of apparent function in the host genome. Transposable elements account for a large proportion of eukaryotic genomes, 45% in humans (LANDER *et al.* 2001), 37.5% in mice (WATERSTON *et al.* 2002), 9% in flies (SPRADLING and RUBIN 1981), 10% in Arabidopsis (INITIATIVE 2000) and even 3% of the diminutive *Saccharomyces* genome (GARFINKEL *et al.* 2005). These numbers fail to account for extremely ancient transposons which have undergone enough substitutions to become unrecognizable, so our understanding of the true contribution of transposons to a genome is partial at best. Researchers have long sought to understand how this genomic bloating occurred, and

what functions it may have. Due to the ancient integration of transposable elements into their host genomes, it is nearly impossible to determine what the organism was like without that transposon, and what changes may have occurred through the incorporation of the transposon into the genome. Do transposons come into a genome as parasites, or do they confer benefits to the organism from the outset? How does a transposon become fixed within a population? These questions are very difficult to answer from a merely historical perspective, and genome forensics can only tell us so much. There are, however, new transposon invasions occurring at the present, and by studying these transposon and host interactions we can finally discard the assumption of transposons as selfish junk, and acquire a better understanding of their role in genome evolution. By studying a recent addition to the genome of *Drosophila melanogaster*, the *P* element, answers to some of these long standing questions may finally be within our reach.

The discovery of the dissociator transposon in maize by Barbara McClintock showed that genetic elements may have the ability to change locations in the genome and also to modulate gene expression (MCCLINTOCK 1950). As the molecular nature of transposons in various organisms was described it became clear that there were two major classes of transposons. The categorization is based on the mechanism of transposition, the Type I transposons move through an RNA intermediate while the Type II transposons, of which the *P* element is a member, move directly as a DNA sequence. Type I transposons are generally more prevalent in genomes as their mechanism of transposition does not alter the parent element. Type II transposons move through a cut-and-paste mechanism which does not directly create new copies of the element. This limitation of proliferation of Type II elements makes successful invasion into new

genomes less dependent on quantity of transpositions, but may be more reliant on the quality of the transpositions. The study of how the newly acquired *P* element in *Drosophila melanogaster* may produce quality transpositions by observing interactions at the *vg* locus may assist in our understanding how Type II elements become successful within a genome.

P elements

P element discovery

In the 1970's it was becoming clear that some strains of *Drosophila melanogaster* appeared to react in an incompatible fashion with other strains. By crossing various lab and wild strains it was found that in certain combinations a syndrome of related phenotypes was observed, which included male crossing over, pupal lethality, sterility, transmission ratio distortion in hybrids, increased female crossing over and chromosomal aberrations (KIDWELL *et al.* 1977). These phenotypes, although not all present in every combination, were linked by a common factor and the collection of these phenotypes was termed hybrid dysgenesis. In addition to the increased rates of chromosome breakage and mutation, the crosses were shown to only have an effect in one direction, that is to say, the reciprocal cross showed no aberrations whatsoever. For this reason the stocks were put into two classes. Those which showed an effect when the strain used a male for the cross were called paternal, or P strains and those which showed an effect when the strain that furnished the female in the cross showing an effect were maternal, or M strains. Using this notation a cross between a P male and an M female would produce an F1 showing hybrid dysgenesis but any other combination of P x P, M x M or M male x P

female would yield normal progeny. A breakthrough in determining the property which gave these strains their mutator potential was the discovery that a cloned 3 kbp piece of DNA from a P strain was able to integrate itself into the genome of M strain flies when microinjected into their developing germline cells (SPRADLING and RUBIN 1982). The ability of these elements to transpose in M strains, but not when injected into P strain organisms, led to the hypothesis that the 3 kbp element encoded a transposase analogous to prokaryotic transposable elements. To account for its behavior in P strains, it was hypothesized that the element also encoded a repressor which accumulated in P strain cells to prohibit further transposition (RUBIN and SPRADLING 1982). These results led Rubin and Spradling to perform the first *P* element transformation by cloning the gene *rosy* into an internally deleted *P* element and injecting a cocktail of this plasmid together with full length *P* element plasmids into *Drosophila* embryos (RUBIN and SPRADLING 1982). Not only did they produce injected flies with a rescued *rosy*$^{+}$ phenotype, but they also produced heritable transformants and thereby revolutionized the study of gene function in *Drosophila.*

_P_ element molecular studies

Sequencing of multiple *P* elements yielded several transposons of differing lengths, but all with conserved 5' and 3' ends (O'HARE and RUBIN 1983). A comparison of these full and internally deleted *P* elements showed a common 31 bp inverted terminal repeat and also an 11 bp inverted internal repeat located within 150 bp of the ends (Figure I-1B) (O'HARE and RUBIN 1983); this study also showed that two-thirds of *P* elements in the average P cytotype genome have large internal deletions. Although the *P* element

contains two repeat regions at each end, the spacing between the 31 bp and 11 bp repeats is unique, there are 94 bp between the repeats at the 5' end and 103 bp between the repeats at the 3' end. This difference is significant as elements which are constructed using two 5' ends or two 3' ends are very poor targets for the transposase (BEALL and RIO 1997; MULLINS *et al.* 1989). Study of the flanking genomic sequence demonstrated a direct 8 bp target site duplication common to all *P* element insertions (O'HARE and RUBIN 1983). Structural and functional analysis of the *P* element by mutagenesis led to the discovery that the four *P* element open reading frames were all required for transposase function. They could not be complemented by additional elements, so the *P* element transposase was a single protein, encoded within the four exons of the gene (KARESS and RUBIN 1984). Transcript analysis showed that there are 2 alternatively spliced forms of the *P* element encoded mRNA (LASKI *et al.* 1986) which would produce the 97 kD transposase and also a 66 kD truncated form of the protein (Figure I-1A, I-1C) (RIO *et al.* 1986). The full length transcript could only be detected in germline, and not somatic cells, but by removal of the third intron a *P* element was produced which encoded full length protein in all tissues (LASKI *et al.* 1986). This construct, commonly referred to as Δ2-3, was found to make transposase in somatic cells showing the requirement of proper splicing for transposase activity. One seemingly minor observation in these experiments, but also worth noting, is that while a polyadenylation site found within the third intron is used for the somatic transcript, there is no polyadenylation site after the fourth exon. Somatic transcripts can be detected that read through the *P* element into the adjacent gene until they encounter a termination site. Further investigation into the tissue specific splicing of the *P* element RNA has shown proper splicing in foreign systems such as frog

oocytes and human cell extracts suggesting that splicing is actively blocked in *Drosophila* somatic cells, rather than having unique splicing factors in the germline (LASKI and RUBIN 1989). The domain required for proper splicing has been further defined as a 20 nucleotide sequence near the 3' end of exon 2 that has been found to bind 50 kD and 97 kD proteins. Mutation of this region produces transcripts which will splice out the third intron in both somatic and germline cells (CHAIN *et al.* 1991). The protein complex that binds this splice blocking site is composed of the host encoded 97 kD *P* element somatic inhibitor (PSI) and the 50 kD hrp48 (SIEBEL *et al.* 1994). The level of PSI is tissue specific with high levels in somatic cells to actively block splicing and low levels in the germline which are insufficient to prevent proper splicing (ADAMS *et al.* 1997). By this tissue specificity PSI has been shown to function as the regulator of *P* element intron 3 splicing.

Repressors and the P cytotype

Although transposase can be produced in germline cells, it does not do so in animals with a so-called P cytotype. The stability of *P* elements in P strains, and the prevention of transposition events in a P female x M male cross is a result of the P cytotype. The control of transposition could be accomplished through one of several mechanisms collectively called P repression. Repression studies first focused on the 66 kD truncated protein to determine what properties it might have. Transformation using elements which could only express the 66 kD protein showed inhibition of a Δ2-3 transposase producing element or an immobile *P* element which contained a wild type transposase gene to mobilize a *P* element carrying a *white*$^{+}$ marker in the germline

(MISRA and RIO 1990). Another result of this study showed that the absolute level of repressor mRNA and protein in the whole organism did not correlate with the level of germline repression, implicating that the spatial and temporal variations of expression from each repressor-producing insert tested were better indicators of repressor ability. In some strains, a large proportion of the mRNA encoding the 66 kD protein was initiated upstream of the *P* element by genomic read-through (MISRA and RIO 1990), indicating that tissue-specific expression can be produced by upstream genomic promoters. Since a large proportion of *P* elements in P strains are internally deleted, Gloor and his co-workers sought to determine the limit to which the 66 kD protein could be truncated and still encode repressor (GLOOR *et al.* 1993). By using 2 different somatic tests of repression, it was shown that proteins which terminated in the first or second exon would not produce repressor. Using the somatic 66 kD transcript as a start, they constructed many 3' end deletions to make further truncated proteins and found that repressor activity was lost in all *P* elements that were truncated 5' to base 1956 (Figure I-1D). These authors designated repressors as Type I and Type II to separate the two distinct classes of *P* elements which showed repressor activity. Those repressors which contained *P* element sequence complete at the 5' end up to base 1956 were classified as Type I repressors while naturally occurring elements with large internal deletions which could also act as repressors were grouped as Type II repressors.

In wild populations the most common Type II repressor is the 1154 bp KP element, first isolated from the Russian strain Krasnodar (BLACK *et al.* 1987). The KP element contains a large deletion from base 808 to 2560 and encodes a 207 aa protein (Figure I-1E) (BLACK *et al.* 1987). Unlike Type I elements, which will only suppress if

maternally inherited, KPs can repress transposition when passed from either parent. The abundance of KP elements in wild strains led to the proposal that KP elements are selected for their ability to prevent *P* element transposition (BLACK *et al.* 1987). Similarly sized elements have also been described with large deletions from the second to fourth exon which have shown varying levels of repressor activity (RASMUSSON *et al.* 1993; SEARLES *et al.* 1986). The two unifying aspects which appear to distinguish Type II elements are zygotic repression irrespective of parent of origin and large internal deletions. However, the range of Type II coding possibilities has not been thoroughly tested.

The mechanism of action for both Type I and Type II repressors is likely similar as they contain the same defined protein motifs. There are at least three different steps in the transposition pathway which may be limited by a repressor protein. The first mechanism involves the repressor protein binding to the full length *P* element and preventing *P* transcription (LEMAITRE and COEN 1991; LEMAITRE *et al.* 1993). In the presence of a functional transposase protein, repressor may also act by competing for the same binding sites on *P* element DNA (KAUFMAN *et al.* 1989; LEE *et al.* 1996). This blockage will prevent the transposase from excising genomic *P* elements. The third mechanism of repression of transposition is multimer poisoning, the presence of dimerization motifs in the repressor and transposase proteins can prevent formation of functional transposase dimers (ANDREWS and GLOOR 1995; LEE *et al.* 1996). There have been some additional studies which have concluded that there are some *P* element insertions which repress transposition in the absence of sequences thought to encode repressor protein (ROCHE and RIO 1998; RONSSERAY *et al.* 2001; RONSSERAY *et al.*

1991). The common feature of these elements is not their sequence, but the location into which they have inserted. The insertion of highly truncated *P* elements into subtelomeric heterochromatin allows for not only the genes within the subtelomeric element to be repressed, but the repression is able to work in trans to shut down *P* element expression genome-wide (RONSSERAY *et al.* 2001). This repression may act through an RNA intermediate, and thus may allow for RNA interference to be an additional mechanism of *P* element repression.

Transposition

Transposition of *P* elements is very amenable to study due to the availability of strains without *P* elements and the ability to study transposition in a cell free system with crude transposase preparations (KAUFMAN and RIO 1992). The *P* elements transpose by a cut-and-paste mechanism where the element is removed from the chromosome leaving a double strand break, and the element is then integrated into a new location while producing the 8 bp target site duplication (KAUFMAN and RIO 1992). The transposase cuts the *P* element, producing a 17 bp 3' overhang at the *P* element ends (BEALL and RIO 1997). It has been found additionally that both a 5' and a 3' end of a *P* element are required for transposition as the enzyme does not readily cut 5' or 3' ends in isolation. This specificity is likely due to the transposase requiring dimerization for proper activity, and the structure of this dimer requires the unique features from the 5' and 3' ends to properly orient itself on the DNA to cut the consensus sequence (BEALL and RIO 1997). The fate of the double strand break created by transposition varies depending on the template used for repair. Transposition can result in the net gain of one *P* element if a

sister chromatid or homologous chromosome is used to recopy the mobilized element, or it can result in the loss of a *P* element from that location if the repair utilizes a wild type copy of the chromosome as in a heterozygous fly (ENGELS *et al.* 1990; GLOOR *et al.* 1991). The double strand break which is created can also be repaired by non-homologous end joining where the resulting chromosome retains approximately the outer half of each 31bp repeat (STAVELEY *et al.* 1995). Integration into a new genomic location is non-random with preference given to transcriptionally active chromatin (BERG and SPRADLING 1991; O'HARE and RUBIN 1983). A more thorough examination of several thousand insertion sites has found there is no base sequence conservation, but there is a common structural element to insertion sites (LIAO *et al.* 2000). Insertion sites are favoured if they have a high GC content in addition to several other computationally defined structural elements. Furthermore, the analysis has revealed a conserved 14 bp structural sequence exists that allows hydrogen bonding between the major groove and the transposase protein. As many different sequences can give similar structures, and the structures are also dependant on other DNA:protein interactions around the target site, it is very difficult to predict whether a given sequence will be favored for insertion based on sequence alone.

Mutagenesis

The insertion of *P* elements into transcriptionally active regions creates the potential for gene mutation. The extent of gene impairment can be affected by insertion site, *P* element sequence and insertion direction. If a *P* element should insert into the coding-region of a gene it will likely induce a frame-shift or premature stop codon

mutation. Insertions into introns may still be detrimental if they introduce a new splice site to include *P* element sequence into the mRNA of the gene. The third location type is into the 5' or 3' untranslated regions which may lead to impairment through loss of mRNA stability, lower translational efficiency or aberrant mRNA localization. Studies of specific *vestigial* alleles have shown that identical insertion sites and insertion orientation can produce varied mutant phenotypes when different internally deleted *P* elements are present (HODGETTS and O'KEEFE 2001; STAVELEY *et al.* 1995; WILLIAMS *et al.* 1988a). Similarly, a careful study of *P* element insertions into the 5' untranslated region of *yellow* has produced 2 alleles with the same insertion site and *P* element sequence, but while one is transcribed collinearly with *yellow*, the other allele transcribes upstream towards the *yellow* promoter (GEYER *et al.* 1991; GEYER *et al.* 1988). These observations of *P* element derivatives of *vestigial* and *yellow* alleles have shown that gene function can vary considerably while retaining the same insertion site. These alleles are more informative as only when certain factors are kept constant between alleles can the nature of phenotypic differences be examined.

Protein and DNA interactions

Comparing the activity and binding affinities of proteins produced by full and internally deleted *P* elements has shed light on what is required for transposition or repression. DNA footprinting studies done using purified proteins and *P* element sequence have revealed the requirements for the protein to bind DNA while several somatic assays of repressor function have been developed which can be used to test *in vivo* various truncated *P* element proteins. By combining these results, the interaction

between protein and target DNA has been well established. Footprinting has shown that transposase recognizes a 10 bp target sequence from base 53 to 62 at the 5' end and 2858 to 2867 at the 3' end of a *P* element, but does not specifically bind either the 31 bp or 11 bp repeats (Figure I-1A) (KAUFMAN *et al.* 1989). This is contrasted with studies using the KP repressor element which show binding to bases 3 to 27, which is within the 31 bp repeat, bases 46 to 68 encompassing the transposase binding site and also bases 122 to 138 which envelops the 11bp repeat with corresponding binding at bases 2855 to 2871, 2757 to 2775 and 2881 to 2905 at the 3' end (Figure I-1C) (LEE *et al.* 1996). This work also showed that while the KP element can bind these sites with high affinity, loss of a previously described leucine zipper motif (Figure I-1A, I-1C) (RIO 1990) results in a greatly decreased affinity for the 31 bp and 11 bp repeat regions (LEE *et al.* 1996). Using more truncated repressors, the DNA binding domain has been localized to the N-terminal 88 amino acids (LEE *et al.* 1998). This region has since been found to contain a conserved domain recently described as the THAP domain (Figure I-1A, I-1C) (ROUSSIGNE *et al.* 2003). The THAP domain is found in several vertebrate and invertebrate species (HAMMER *et al.* 2005; QUESNEVILLE *et al.* 2005) and is located in the first 83 amino acids of the *P* element, corresponding precisely with the known DNA binding region of the *P* element encoded proteins (ROUSSIGNE *et al.* 2003). The question of whether binding ability corresponds to repression can be answered using somatic tests of repressor activity. A favoured system employs a tandem insertion of *P* elements into the bristle patterning gene, *singed* (*sn*). The weak bristle phenotype of this so-called *singed*^{weak} allele is highly stable in the absence of additional *P* elements. However, supplying somatic transposase causes mobilization of either one, or both, of the *P*

elements from this location and can produce a wild type allele or a severe bristle phenotype that is easily scored depending on the element which is lost (NITASAKA and YAMAZAKI 1988; ROIHA *et al.* 1988). This system requires transposase to assay repressor function but there is a system using the wing patterning gene *vestigial* (*vg*), which can be tested using only repressor. The vg^{21-3} allele causes a strong wing mutation in the absence of repressor, but the phenotype is largely restored to wild type when repressor making elements are added (WILLIAMS *et al.* 1988a). Prior to the work described in this thesis, repressor studies have not shown any proteins shorter than that encoded by the KP element to work as repressors *in vivo*.

Population biology

The presence of strains of *Drosophila melanogaster* in various labs which did not have any *P* elements was the essential factor leading to their discovery, but once it was determined how *P* elements cause hybrid dysgenesis, the question still remained of why the lab populations had no *P* elements. As further population studies revealed that flies in the wild currently all contain *P* elements it was also noted that numerous stocks established before 1950 did not contain any *P* elements (KIDWELL 1983). The two plausible explanations for this difference were that *P* elements were recent genomic invaders, or they were lost in lab stocks over the years they have been cultured. The rapid invasion hypothesis has been supported by the lack of *P* element remnants in M strains, and the near perfect sequence match (2906 of the 2907 bases) between the *D. melanogaster* and *D. willistoni P* elements (DANIELS *et al.* 1990). This recent acquisition of *P* elements into the *Drosophila melanogaster* genome would suggest horizontal

transfer from a related species containing *P* elements. The likely vector of the horizontal transfer is a semiparasitic mite (*Proctolaelaps regalis*), which parasitizes all immature stages of *Drosophila* and has been demonstrated to carry *P* element DNA when grown in the presence of P strains (HOUCK *et al.* 1991). The mites feed on eggs by rapid injection of their needle-like mouthparts, lasting usually less than one second per feed and sometimes followed by immediate feeding on adjacent eggs. The protocol of microinjection of *P* element DNA to produce transgenic flies is strikingly similar to this feeding process (RUBIN and SPRADLING 1982).

The highly deleterious effects of hybrid dysgenesis on fertility and overall fitness in *Drosophila melanogaster* complicate our understanding of how they became fixed in the species throughout the world within a 50 year time frame. Indeed, the majority of *P* element insertions result in lethal alleles or alleles conferring reduced fitness (LYMAN *et al.* 1996; MACKAY 1989; PASYUKOVA *et al.* 2004). By investigating how a *P* element induced deleterious mutation of the wing determining gene *vestigial* (*vg*) can be repressed by the presence of additional *P* elements, this study adds to our understanding of how the rise of *P* elements did not abolish the fitness of the *Drosophila melanogaster* species.

Vestigial – Cloning and *P* element alleles

Vestigial is one of the classic *Drosophila* mutants, first described by Morgan (BRIDGES and MORGAN 1919). The molecular description of a *vg* mutant came 75 years later with the recovery of a cryptic *P* element-induced mutant designated vg^{21} (WILLIAMS and BELL 1988). This allowed the locus to be cloned and molecularly characterized (WILLIAMS *et al.* 1990; WILLIAMS and BELL 1988). The gene is found to span

approximately 16 kbp (WILLIAMS *et al.* 1990) and contains 8 exons (Figure I-2A) (WILLIAMS *et al.* 1991). The 453 aa VG product is encoded by exons 2 through 8 of the *vg* mRNA. More severe *P* element alleles, arising from secondary events involving the vg^{21} insert have been generated in dysgenic crosses, including $vg^{21\text{-}3}$ (WILLIAMS *et al.* 1988a) and vg^{2a33} (HODGETTS and O'KEEFE 2001). All *P* element-induced *vg* alleles have their *P* element inserted in the same place in the 5'-untranslated region of the first exon (Figure I-2a), with the transposon promoter running in parallel with that of the *vestigial* promoter (HODGETTS and O'KEEFE 2001; STAVELEY *et al.* 1994; WILLIAMS *et al.* 1988a). The phenotypic variation seen in these *P* element *vg* alleles must be caused by some aspect of *P* element biology. An unusual feature of the *P* element alleles of *vg* is that the mutant phenotype of some alleles is suppressed in the presence of *P* elements that encode a repressor whereas other alleles are insensitive to P repressor (HODGETTS and O'KEEFE 2001; WILLIAMS *et al.* 1988b). Sensitivity to P repressor, or the lack of it, must be caused by differences in the internal sequences of the *P* elements, since the insertion point and orientation of the *P* element in every allele are identical.

Vestigial function

In order to use the interaction between *vestigial* and P element insertions as a basis of understanding a mechanism of *P* element mutation, we need to also understand the role of *vestigial* in a developing fly. In the hierarchy of genes required to produce the fly wing, *vestigial* is at the top with another gene *scalloped*, which function as wing selector genes (HALDER *et al.* 1998; WILLIAMS *et al.* 1993). Together these 2 proteins form a transcriptional activator complex (SIMMONDS *et al.* 1998) which reinforces both

the expression of each other (HALDER *et al.* 1998) and turns on downstream wing genes (GUSS *et al.* 2001). The expression of *vg* can be broken down into three spatial and temporal domains. Early *vg* expression is seen ubiquitously in the embryonic wing primordia and is one of the earliest markers of wing fate (WILLIAMS *et al.* 1991; WILLIAMS *et al.* 1993). The differential patterning of wing regions by *vestigial* expression occurs during the third larval instar and is the result of two tissue specific enhancers. The boundary enhancer, *vg*(BE) (Figure I-2A), is located within the second intron and drives the expression of VG along the dorsal-ventral compartmental boundary of the wing disc (WILLIAMS *et al.* 1994), which will become the margin of the adult wing (Figure I-2B). The *vg*(BE) drives expression in response to the Notch pathway (KIM *et al.* 1995) and is the first localized expression of VG in the wing. The later acting tissue specific enhancer is the quadrant enhancer, *vg*(QE) (Figure I-2A), which is located in the fourth intron (KIM *et al.* 1996). The *vg*(QE) causes the expression of *vestigial* by *decapentaplegic* in the four quadrants of the circular wing pouch as divided by the AP and DV boundaries (Figure I-2C) (KIM *et al.* 1996). Once the proper expression of *vestigial* is induced, several downstream genes such as *cut, spalt, homothorax, extradenticle* and *wingless* are controlled to produce wing tissue (AZPIAZU and MORATA 2000; GUSS *et al.* 2001).

The *vg – P* element system

As the wing is a complex structure with defined landmarks such as size, shape and location of veins, it is possible to very clearly identify aberrations in the proper pattern. *P* element alleles of *vestigial* exist which show a gradation from full wings to

nicked wings, strap wings, small stump-like wings and finally the complete absence of wings. This graded loss of wing tissue has been used to produce a quantitative scale for measuring the degree of malformation of the wing (NAKASHIMA-TANAKA 1968a; NAKASHIMA-TANAKA 1968b). This scale of 31 wing phenotypes is in some cases grouped into a 6 point scale which can be used to categorize a mutant by average severity, rather than by the range of phenotypes seen (Figure I-3) (WILLIAMS *et al.* 1988a). The sensitivity with which we can detect *vestigial* function by a measure of wing morphology produces an excellent system to study the effect of *P* element insertions at *vestigial* and also transposon and genomic modifiers of the phenotype. By determining some of the interplay between transposon and genome at this non-essential, but surely beneficial gene, a better understanding of the role of transposons may be attained.

Figure I-1: Sequence elements within the *P* element transposon. The middle line **(B)** shows the 31 bp repeats in dark yellow and 11 bp repeats in orange over the full length *P* element. The exons are marked along the sequence in pink. Above the element **(A)** the open reading frame in the germ line spliced mRNA which gives rise to the transposase is denoted with the two sites the protein binds marked in red. Below the element **(C)** is the open reading frame present in the somatic transcript which does not splice out the third intron. The protein produced from this mRNA will bind to six sites along the *P* element sequence which are highlighted in pink. The protein domain responsible for DNA binding (THAP domain) is marked as a dark grey box with the text "THAP" on the open reading frames in **(A)** and **(C)**. The leucine zipper dimerization domain is marked on the open reading frames of **(A)** and **(C)** with a light grey box labeled "LZ". The minimum sequence required to produce a Type I repressor is marked in **(D)**, while the sequence present in the most common Type II repressor making element, the KP element, is marked in **(E)**.

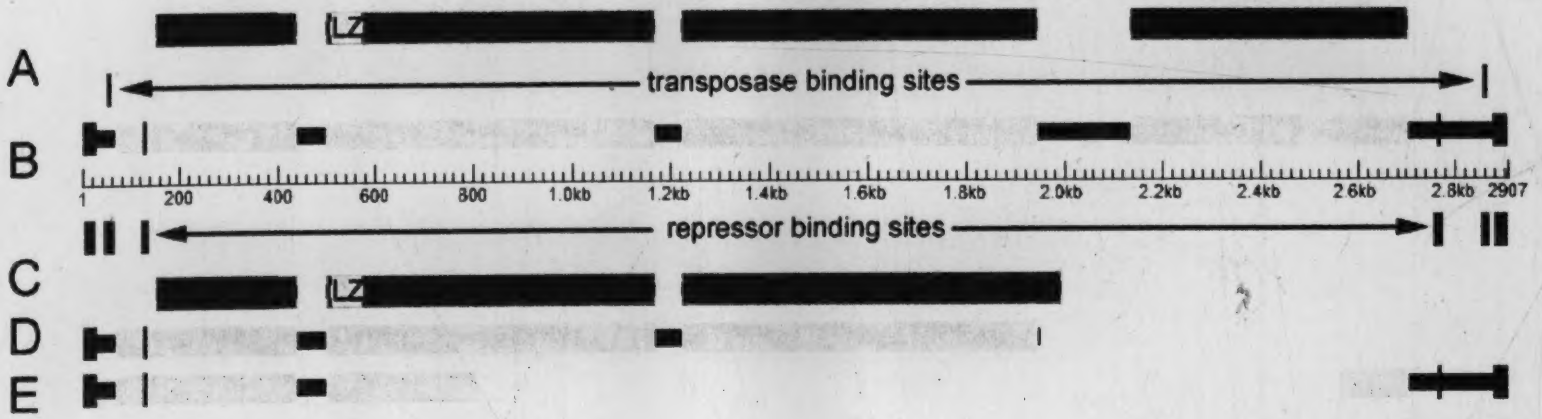

A
B
C
D
E
LZ
LZ
transposase binding sites
repressor binding sites
1
200
400
600
800
1.0kb
1.2kb
1.4kb
1.6kb
1.8kb
2.0kb
2.2kb
2.4kb
2.6kb
2.8kb
2907

Figure I-2: Sequence elements within the *vg* gene. (A) The eight exons of the *vestigial* gene are marked in blue, with the quadrant enhancer vg(QE) and boundary enhancer vg(BE) also noted. The insertion point of all *P* element *vestigial* alleles is designated by the black triangle above the first exon. The protein coding region of the *vg* transcript is shown in green above the corresponding genomic sequence. **(B,C)** The region of VG product expression driven by each enhancer in a third instar wing disc is shown for the vg(BE) and vg(QE) respectively.

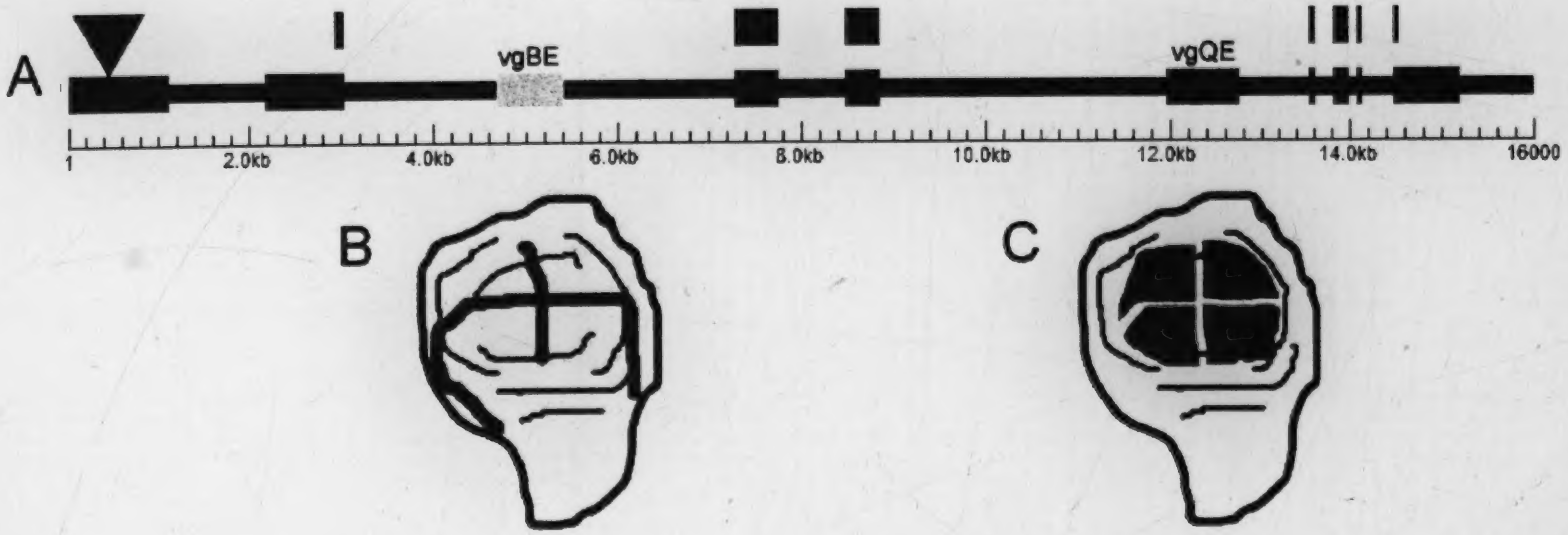
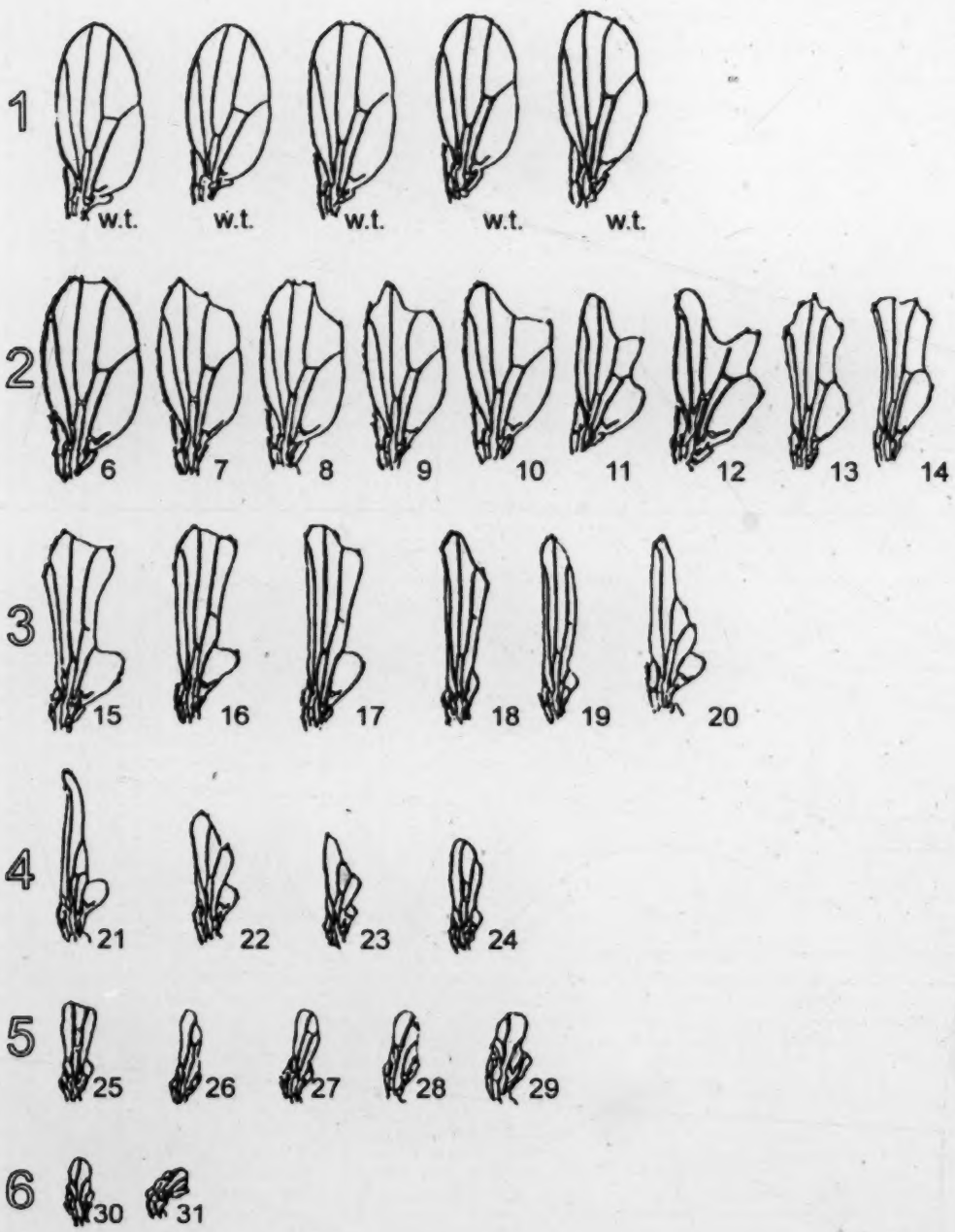

A
vgBE
vgQE
1
2.0kb
4.0kb
6.0kb
8.0kb
10.0kb
12.0kb
14.0kb
16000
B
C

Figure I-3: Wing phenotype severity scale. The *vestigial* mutant wing scale using original wing drawings as published by Nakashima-Tanaka, 1968. The two scales used in this study are the 6 point generalized scale (outlined numbers) initially used by Jim Williams to classify *P* element *vestigial* alleles, and a modification of the 31 point scale (solid numbers) where numbers 1 through 5 are replaced by the notation w.t., meaning wild type, to denote an absence of any discontinuities in the wing margin.


1
w.t.
w.t.
w.t.
w.t.
w.t.
2
6
7
8
9
10
11
12
13
14
3
15
16
17
18
19
20
4
21
22
23
24
5
25
26
27
28
29
6
30
31

Literature Cited

ADAMS, M. D., R. S. TARNG and D. C. RIO, 1997 The alternative splicing factor PSI regulates P-element third intron splicing in vivo. Genes Dev **11:** 129-138.

ANDREWS, J. D., and G. B. GLOOR, 1995 A role for the KP leucine zipper in regulating P element transposition in Drosophila melanogaster. Genetics **141:** 587-594.

AZPIAZU, N., and G. MORATA, 2000 Function and regulation of homothorax in the wing imaginal disc of Drosophila. Development **127:** 2685-2693.

BEALL, E. L., and D. C. RIO, 1997 Drosophila P-element transposase is a novel site-specific endonuclease. Genes Dev **11:** 2137-2151.

BERG, C. A., and A. C. SPRADLING, 1991 Studies on the rate and site-specificity of P element transposition. Genetics **127:** 515-524.

BLACK, D. M., M. S. JACKSON, M. G. KIDWELL and G. A. DOVER, 1987 KP elements repress P-induced hybrid dysgenesis in Drosophila melanogaster. Embo J **6:** 4125-4135.

BRIDGES, C. B., and T. H. MORGAN, 1919 Contributions to the genetics of Drosophila melanogaster. Publs Carnegie Instn **278:** 123-304.

CHAIN, A. C., S. ZOLLMAN, J. C. TSENG and F. A. LASKI, 1991 Identification of a cis-acting sequence required for germ line-specific splicing of the P element ORF2-ORF3 intron. Mol Cell Biol **11:** 1538-1546.

DANIELS, S. B., K. R. PETERSON, L. D. STRAUSBAUGH, M. G. KIDWELL and A. CHOVNICK, 1990 Evidence for horizontal transmission of the P transposable element between Drosophila species. Genetics **124:** 339-355.

ENGELS, W. R., D. M. JOHNSON-SCHLITZ, W. B. EGGLESTON and J. SVED, 1990 High-frequency P element loss in Drosophila is homolog dependent. Cell **62:** 515-525.

GARFINKEL, D. J., K. M. NYSWANER, K. M. STEFANISKO, C. CHANG and S. P. MOORE, 2005 Ty1 copy number dynamics in Saccharomyces. Genetics **169:** 1845-1857.

GEYER, P. K., A. J. CHIEN, V. G. CORCES and M. M. GREEN, 1991 Mutations in the su(s) gene affect RNA processing in Drosophila melanogaster. Proc Natl Acad Sci U S A **88:** 7116-7120.

GEYER, P. K., K. L. RICHARDSON, V. G. CORCES and M. M. GREEN, 1988 Genetic instability in Drosophila melanogaster: P-element mutagenesis by gene conversion. Proc Natl Acad Sci U S A **85:** 6455-6459.

GLOOR, G. B., N. A. NASSIF, D. M. JOHNSON-SCHLITZ, C. R. PRESTON and W. R. ENGELS, 1991 Targeted gene replacement in Drosophila via P element-induced gap repair. Science **253:** 1110-1117.

GLOOR, G. B., C. R. PRESTON, D. M. JOHNSON-SCHLITZ, N. A. NASSIF, R. W. PHILLIS *et al.*, 1993 Type I repressors of P element mobility. Genetics **135:** 81-95.

GUSS, K. A., C. E. NELSON, A. HUDSON, M. E. KRAUS and S. B. CARROLL, 2001 Control of a genetic regulatory network by a selector gene. Science **292:** 1164-1167.

HALDER, G., P. POLACZYK, M. E. KRAUS, A. HUDSON, J. KIM *et al.*, 1998 The Vestigial and Scalloped proteins act together to directly regulate wing-specific gene expression in Drosophila. Genes Dev **12:** 3900-3909.

HAMMER, S. E., S. STREHL and S. HAGEMANN, 2005 Homologs of Drosophila P transposons were mobile in zebrafish but have been domesticated in a common ancestor of chicken and human. Mol Biol Evol **22:** 833-844.

HODGETTS, R. B., and S. L. O'KEEFE, 2001 The mutant phenotype associated with P-element alleles of the vestigial locus in Drosophila melanogaster may be caused by a readthrough transcript initiated at the P-element promoter. Genetics **157:** 1665-1672.

HOUCK, M. A., J. B. CLARK, K. R. PETERSON and M. G. KIDWELL, 1991 Possible horizontal transfer of Drosophila genes by the mite Proctolaelaps regalis. Science **253:** 1125-1128.

INITIATIVE, T. A. G., 2000 Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature **408:** 796-815.

KARESS, R. E., and G. M. RUBIN, 1984 Analysis of P transposable element functions in Drosophila. Cell **38:** 135-146.

KAUFMAN, P. D., R. F. DOLL and D. C. RIO, 1989 Drosophila P element transposase recognizes internal P element DNA sequences. Cell **59:** 359-371.

KAUFMAN, P. D., and D. C. RIO, 1992 P element transposition in vitro proceeds by a cut-and-paste mechanism and uses GTP as a cofactor. Cell **69:** 27-39.

KIDWELL, M. G., 1983 Evolution of hybrid dysgenesis determinants in Drosophila melanogaster. Proc Natl Acad Sci U S A **80:** 1655-1659.

KIDWELL, M. G., J. F. KIDWELL and J. A. SVED, 1977 Hybrid dysgenesis in Drosophila melanogaster: a syndrome of aberrant traits including mutation, sterility and male recombination. Genetics **86:** 813-833.

KIM, J., K. D. IRVINE and S. B. CARROLL, 1995 Cell recognition, signal induction, and symmetrical gene activation at the dorsal-ventral boundary of the developing Drosophila wing. Cell **82:** 795-802.

KIM, J., A. SEBRING, J. J. ESCH, M. E. KRAUS, K. VORWERK *et al.*, 1996 Integration of positional signals and regulation of wing formation and identity by Drosophila vestigial gene. Nature **382:** 133-138.

LANDER, E. S., L. M. LINTON, B. BIRREN, C. NUSBAUM, M. C. ZODY *et al.*, 2001 Initial sequencing and analysis of the human genome. Nature **409:** 860-921.

LASKI, F. A., D. C. RIO and G. M. RUBIN, 1986 Tissue specificity of Drosophila P element transposition is regulated at the level of mRNA splicing. Cell **44:** 7-19.

LASKI, F. A., and G. M. RUBIN, 1989 Analysis of the cis-acting requirements for germ-line-specific splicing of the P-element ORF2-ORF3 intron. Genes Dev **3:** 720-728.

LEE, C. C., E. L. BEALL and D. C. RIO, 1998 DNA binding by the KP repressor protein inhibits P-element transposase activity in vitro. Embo J **17:** 4166-4174.

LEE, C. C., Y. M. MUL and D. C. RIO, 1996 The Drosophila P-element KP repressor protein dimerizes and interacts with multiple sites on P-element DNA. Mol Cell Biol **16:** 5616-5622.

LEMAITRE, B., and D. COEN, 1991 P regulatory products repress in vivo the P promoter activity in P-lacZ fusion genes. Proc Natl Acad Sci U S A **88:** 4419-4423.

LEMAITRE, B., S. RONSSERAY and D. COEN, 1993 Maternal repression of the P element promoter in the germline of Drosophila melanogaster: a model for the P cytotype. Genetics **135:** 149-160.

LIAO, G. C., E. J. REHM and G. M. RUBIN, 2000 Insertion site preferences of the P transposable element in Drosophila melanogaster. Proc Natl Acad Sci U S A **97:** 3347-3351.

LYMAN, R. F., F. LAWRENCE, S. V. NUZHDIN and T. F. MACKAY, 1996 Effects of single P-element insertions on bristle number and viability in Drosophila melanogaster. Genetics **143:** 277-292.

MACKAY, T. F., 1989 Transposable elements and fitness in Drosophila melanogaster. Genome **31:** 284-295.

MCCLINTOCK, B., 1950 The origin and behavior of mutable loci in maize. Proc Natl Acad Sci U S A **36:** 344-355.

MISRA, S., and D. C. RIO, 1990 Cytotype control of Drosophila P element transposition: the 66 kd protein is a repressor of transposase activity. Cell **62:** 269-284.

MULLINS, M. C., D. C. RIO and G. M. RUBIN, 1989 cis-acting DNA sequence requirements for P-element transposition. Genes Dev **3:** 729-738.

NAKASHIMA-TANAKA, E., 1968a The effect of temperature and genetic background on the phenotypic expression of several vestigial strains of Drosophila melanogaster. Genetica **38:** 447-458.

NAKASHIMA-TANAKA, E., 1968b The effects of chemicals on the phenotypic expression of a vestigial mutant in Drosophila melanogaster. Genetica **38:** 459-470.

NITASAKA, E., and T. YAMAZAKI, 1988 Excision of one of two defective P elements as the cause of alternate mutational events (sn+ and sne) of the singed bristle allele snw in Drosophila melanogaster. Jpn J Genet **63:** 303-312.

O'HARE, K., and G. M. RUBIN, 1983 Structures of P transposable elements and their sites of insertion and excision in the Drosophila melanogaster genome. Cell **34:** 25-35.

PASYUKOVA, E. G., S. V. NUZHDIN, T. V. MOROZOVA and T. F. MACKAY, 2004 Accumulation of transposable elements in the genome of Drosophila melanogaster is associated with a decrease in fitness. J Hered **95:** 284-290.

QUESNEVILLE, H., D. NOUAUD and D. ANXOLABEHERE, 2005 Recurrent recruitment of the THAP DNA-binding domain and molecular domestication of the P-transposable element. Mol Biol Evol **22:** 741-746.

RASMUSSON, K. E., J. D. RAYMOND and M. J. SIMMONS, 1993 Repression of hybrid dysgenesis in Drosophila melanogaster by individual naturally occurring P elements. Genetics **133:** 605-622.

RIO, D. C., 1990 Molecular mechanisms regulating Drosophila P element transposition. Annu Rev Genet **24:** 543-578.

RIO, D. C., F. A. LASKI and G. M. RUBIN, 1986 Identification and immunochemical analysis of biologically active Drosophila P element transposase. Cell **44:** 21-32.

ROCHE, S. E., and D. C. RIO, 1998 Trans-silencing by P elements inserted in subtelomeric heterochromatin involves the Drosophila Polycomb group gene, Enhancer of zeste. Genetics **149:** 1839-1855.

ROIHA, H., G. M. RUBIN and K. O'HARE, 1988 P element insertions and rearrangements at the singed locus of Drosophila melanogaster. Genetics **119:** 75-83.

RONSSERAY, S., A. BOIVIN and D. ANXOLABEHERE, 2001 P-Element repression in Drosophila melanogaster by variegating clusters of P-lacZ-white transgenes. Genetics **159:** 1631-1642.

RONSSERAY, S., M. LEHMANN and D. ANXOLABEHERE, 1991 The maternally inherited regulation of P elements in Drosophila melanogaster can be elicited by two P copies at cytological site 1A on the X chromosome. Genetics **129:** 501-512.

ROUSSIGNE, M., S. KOSSIDA, A. C. LAVIGNE, T. CLOUAIRE, V. ECOCHARD *et al.*, 2003 The THAP domain: a novel protein motif with similarity to the DNA-binding domain of P element transposase. Trends Biochem Sci **28:** 66-69.

RUBIN, G. M., and A. C. SPRADLING, 1982 Genetic transformation of Drosophila with transposable element vectors. Science **218:** 348-353.

SEARLES, L. L., A. L. GREENLEAF, W. E. KEMP and R. A. VOELKER, 1986 Sites of P element insertion and structures of P element deletions in the 5' region of Drosophila melanogaster RpII215. Mol Cell Biol **6:** 3312-3319.

SIEBEL, C. W., R. KANAAR and D. C. RIO, 1994 Regulation of tissue-specific P-element pre-mRNA splicing requires the RNA-binding protein PSI. Genes Dev **8:** 1713-1725.

SIMMONDS, A. J., X. LIU, K. H. SOANES, H. M. KRAUSE, K. D. IRVINE *et al.*, 1998 Molecular interactions between Vestigial and Scalloped promote wing formation in Drosophila. Genes Dev **12:** 3815-3820.

SPRADLING, A. C., and G. M. RUBIN, 1981 Drosophila genome organization: conserved and dynamic aspects. Annu Rev Genet **15:** 219-264.

SPRADLING, A. C., and G. M. RUBIN, 1982 Transposition of cloned P elements into Drosophila germ line chromosomes. Science **218:** 341-347.

STAVELEY, B. E., T. R. HESLIP, R. B. HODGETTS and J. B. BELL, 1995 Protected P-element termini suggest a role for inverted-repeat-binding protein in transposase-induced gap repair in Drosophila melanogaster. Genetics **139:** 1321-1329.

STAVELEY, B. E., R. B. HODGETTS, S. L. O'KEEFE and J. B. BELL, 1994 Targeting of an enhancer trap to vestigial. Dev Biol **165:** 290-293.

WATERSTON, R. H., K. LINDBLAD-TOH, E. BIRNEY, J. ROGERS, J. F. ABRIL *et al.*, 2002 Initial sequencing and comparative analysis of the mouse genome. Nature **420:** 520-562.

WILLIAMS, J. A., A. L. ATKIN and J. B. BELL, 1990 The functional organization of the vestigial locus in Drosophila melanogaster. Mol Gen Genet **221:** 8-16.

WILLIAMS, J. A., and J. B. BELL, 1988 Molecular organization of the vestigial region in Drosophila melanogaster. Embo J **7:** 1355-1363.

WILLIAMS, J. A., J. B. BELL and S. B. CARROLL, 1991 Control of Drosophila wing and haltere development by the nuclear vestigial gene product. Genes Dev **5:** 2481-2495.

WILLIAMS, J. A., S. W. PADDOCK and S. B. CARROLL, 1993 Pattern formation in a secondary field: a hierarchy of regulatory genes subdivides the developing Drosophila wing disc into discrete subregions. Development **117:** 571-584.

WILLIAMS, J. A., S. W. PADDOCK, K. VORWERK and S. B. CARROLL, 1994 Organization of wing formation and induction of a wing-patterning gene at the dorsal/ventral compartment boundary. Nature **368:** 299-305.

WILLIAMS, J. A., S. S. PAPPU and J. B. BELL, 1988a Molecular analysis of hybrid dysgenesis-induced derivatives of a P-element allele at the vg locus. Mol Cell Biol **8:** 1489-1497.

WILLIAMS, J. A., S. S. PAPPU and J. B. BELL, 1988b Suppressible P-element alleles of the vestigial locus in Drosophila melanogaster. MGG **212:** 370-374.

Chapter II: Materials and Methods

β-galactosidase staining: Dissected larvae and whole adults were fixed in PBS with 0.75% gluteraldehyde for 20 minutes. After fixation they were washed 3 times in PBT (PBS with 0.05% Triton X-100). The samples were then stained in the detection solution (10 mM Na_2HPO_4, 150 mM NaCl, 1 mM $MgCl_2$, 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 0.2% 5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside, pH 7.2) for 2-12 hours at 37°C. The samples were then washed 3 times in PBT to remove excess detection solution and examined under the microscope.

DNA extraction: Genomic DNA was obtained by placing 250 to 500 μL of flies into a 1.5 mL tube and adding 500-1000 μL of TESS (10 mM Tris-Cl pH 7.5, 60 mM EDTA, 0.15 mM spermine and 0.15 mM spermidine) and 5 mL of 10 mg/mL Proteinase K. The mixture was ground up using a pestle and incubated at 37°C. After 30 minutes 500 to 1000 μL of TES (0.2 M Tris-Cl pH 9.0, 30 mM EDTA, 2% Sodium Dodecyl Sulfate) were added for 30 minutes at 65°C. The solution was then extracted using phenol, then a 1:1 mixture of phenol and chloroform and finally chloroform to remove proteins from the solution. The DNA was precipitated by adding 4 M NaCl to a 0.2 M final concentration and then 2.5 volumes of 95% ethanol. The tubes were placed on ice for 15 minutes and then the DNA was pelleted for 10 minutes at full speed in a microcentrifuge. The pellet was washed using 70% ethanol, dried and then resuspended in 500 μL TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA). To remove any RNA 1 μL of a 10 mg/mL RNAse A solution was added during the reconstitution process. After 15 minutes at 37°C the

solution was once again extracted using a mixture of phenol and chloroform and then again using just chloroform to remove any residual proteins. A second precipitation was done as before and the pellet is suspended in 50 µL TE pH 8.0.

***Drosophila* stocks:** Stocks were maintained at 25°C on a standard cornmeal/molasses medium (96 mL fancy molasses, 72 g cornmeal, 18 g yeast, 8 g agar, 2.8 g methyl paraben in 10 mL ethanol and 2.8 ml propionic acid per liter of food). For a list of stocks used see Table II-1.

Germline transformation: Injections we carried out using a modification of published methods (Rubin and Spradling 1982). Collections of embryos in population cages were done for 45 minutes on 4% agar plates supplemented with 10% concentrated Welch's grape juice and approximately 200 µL yeast paste. Embryos were then dechorionated using 50% bleach for 55 seconds, and then lined up on the same grape agar before transferring to a slide with a coverslip mounted on it that was coated in glue to hold the embryos. Desiccation of the embryos for 8 minutes followed and then the embryos were covered with medium halocarbon oil. A mixture of 5 µg pUChsΔ2–3 and 25 µg of either pGEM-T-vg^{21} or pH-Stinger[21-3] *P* element plasmid was resuspended in 50 µL water and used for injection. After injection, embryos were incubated 1-3 days at 18°C in a humidified chamber and first instar larvae were transferred to softened food until adults emerged. Approximately 400 embryos were injected for each construct with an adult survival rate around 50%. For injection of the pGEM-T-vg^{21} plasmid the host used was *w; cn* $vg^{21\text{-}3}$ and for the pH-Stinger[21-3] plasmid the host used was y^1 w^{1118}.

Inverse PCR: 50 µg of vg^{21r7} genomic DNA were digested with 150 units of *Sau*3AI overnight in a volume of 150 µl. After 1 hour at 65°C the DNA was ligated overnight with 15 U T4 DNA ligase in a total volume of 500 µl to facilitate circularization. The Kan11 and Kan12 primers (see Table II-2) were designed to yield an inverse PCR product of 659 bp plus the distance to the 5' *Sau*3AI site. Standard PCR using Kan11 and Kan12 primers and 2 µl of the ligation mixture yielded several bands which were purified using a QIAquick Gel Extraction Kit (Qiagen) and directly sequenced to determine the location of new *P* element insertions.

PCR: Genomic DNA was amplified using appropriate primers located throughout the vg region and P element sequence (see Figure II-1). Fragments to be further characterized were amplified using a *Taq*:*Pfu* mixture and were cloned into pGEM-T Easy (Promega). For samples which proved difficult to amplify I used High Fidelity Platinum *Taq* (Invitrogen) according to the manufacturer's instructions.

Plasmids: Transposase source pUChsΔ2–3 (Flybase ID FBmc-0000938) was a gift from Shelagh Campbell. The pGEM-T-vg^{21} plasmid was constructed by cloning a 1704 bp PCR product of SOK34 to CHU1 (see Figure II-1) from genomic DNA of vg^{21} flies into pGEM-T Easy (Promega). The pH-Stinger[21-3] plasmid was constructed by a blunt end ligation of a THE5 to THE1 (see Figure II-1) PCR product of *w; cn* $vg^{21\text{-}3}$ placed into the *Bam* HI site of pH-Stinger (a gift from S. Barolo). The resulting plasmids were assayed

for insertional direction and the desired product, in which the internal *P* element was inserted into the opposite orientation of the vector *P* element, was chosen.

RNA extraction: RNA was extracted from various developmental stages by homogenization in Trizol reagent (Invitrogen) according to the manufacturer's instructions.

S1 analysis: RNA protection assays were done using a Multi-NPA kit (Ambion). Pupal RNA was extracted, DNAse I treated and resuspended in 50 μL deionized formamide. After quantification, 50 μg of each sample was mixed with an end-labeled, polyacrylamide gel size-purified, radioactive oligonucleotide probe according to the manufacturer's instructions and hybridized at 34°C overnight. S1 nuclease digestion followed hybridization using the manufacturer's instructions. The optimal nuclease digestion mixture was found to be 1:200 dilution of nuclease used at 26°C for 30 minutes. After digestion and precipitation the samples were run on a denaturing polyacrylamide gel (12% polyacrylamide [40:1 ratio of acrylamide to bisacrylamide]), 7 M urea, 0.5x TBE (45 mM Tris, 45 mM boric acid, 1 mM EDTA) to resolve the bands. The products were visualized using either film or a Molecular Dynamics Phosphorimager.

Selection of new *vg* alleles: Flies from the microinjections were crossed to vg^{B}/CyO flies and the resulting progeny were scored for wing phenotype. Flies which showed a less severe wing morphology than $vg^{21\text{-}3}/vg^{B}$ were selected as potential new *vg* alleles. The vg^{*}/vg^{B} flies were then crossed to vg^{B}/CyO and the *CyO* flies were individually crossed to

vg^{B}/*CyO* one last time. The crosses which yielded only *CyO* flies were discarded as carrying vg^{B} and from the other crosses, the vg^{*}/*CyO* were sib-crossed to yield homozygous stocks of the new alleles.

Sequencing: Templates were sequenced using DYEnamic ET terminators and run on an ABI 373 automated sequencer (Amersham Biosciences, Piscataway, NJ) according to manufacturer's instructions using appropriately located PCR primers. All sequencing was done for both forward and reverse strands with discrepancies resequenced to find a consensus. Sequencing reactions were processed by the Department of Biological Sciences Molecular Biology Service Unit (MBSU).

Slot blots: The slot blot apparatus was washed with 3% hydrogen peroxide for 30 minutes to remove any RNAses and then washed several times with water. Unused slots were filled with a 3% agarose solution to prevent solutions getting suctioned into those areas. Samples were 25 μg of total RNA resuspended in 500 μl 10 mM NaOH, 1 mM EDTA. The samples were drawn through a positively charged nylon membrane (GeneScreen Plus; DuPont) and then the slots were washed with an additional 1 ml of 10 mM NaOH, 1 mM EDTA. The membrane was then hybridized according to manufacturer's instructions with an *RpL32* probe (produced using a genomic *RpL32* PCR product, α^{32}P-dCTP and a DECAprime II (Ambion) random priming DNA labeling kit) to allow quantification of RNA levels across samples. The products were visualized using either film or a Molecular Dynamics Phosphorimager.

Southern analysis: Approximately 5 μg of genomic DNA was digested and transferred to positively charged nylon membranes (GeneScreen Plus; DuPont) using a downward transfer in 0.4 M NaOH (KOETSIER *et al.* 1993). Membranes were then hybridized with radioactive probes according to the manufacturer's instructions. Probes were made using dsDNA fragments which were randomly labeled using a DECAprime II kit (Ambion). Approximately $1.5x10^6$ cpm of probe/mL hybridization buffer was used.

Wing quantification: Wings were dissected, mounted onto slides and photographed using a Pentax Optio-S 3.1 megapixel digital camera. The images were opened in Adobe Photoshop CS1 and the wings were traced and the pixel area was reported using the histogram function. The pixel area was then converted to actual area by using a picture of a stage micrometer taken at the same time as the wing pictures to determine the number of pixels per μm.

Figure II-1: *P* element and *vg* PCR primer locations. Exons of the *P* element are shown in pink and the first two exons of *vestigial* are shown in blue with the *P* element insertion site shown in yellow. **(A)** Primers used for PCR of the *vg* region are shown along the 5' end of the *vg* gene. Forward primers are shown in green, while reverse primers are shown in red. The *P* element insertion site is shown as a yellow line within the first *vg* exon. **(B)** Primers used for PCR of *P* element sequences are shown with forward primers in green and reverse primers in red. All primer sequences are listed in Table II-2.

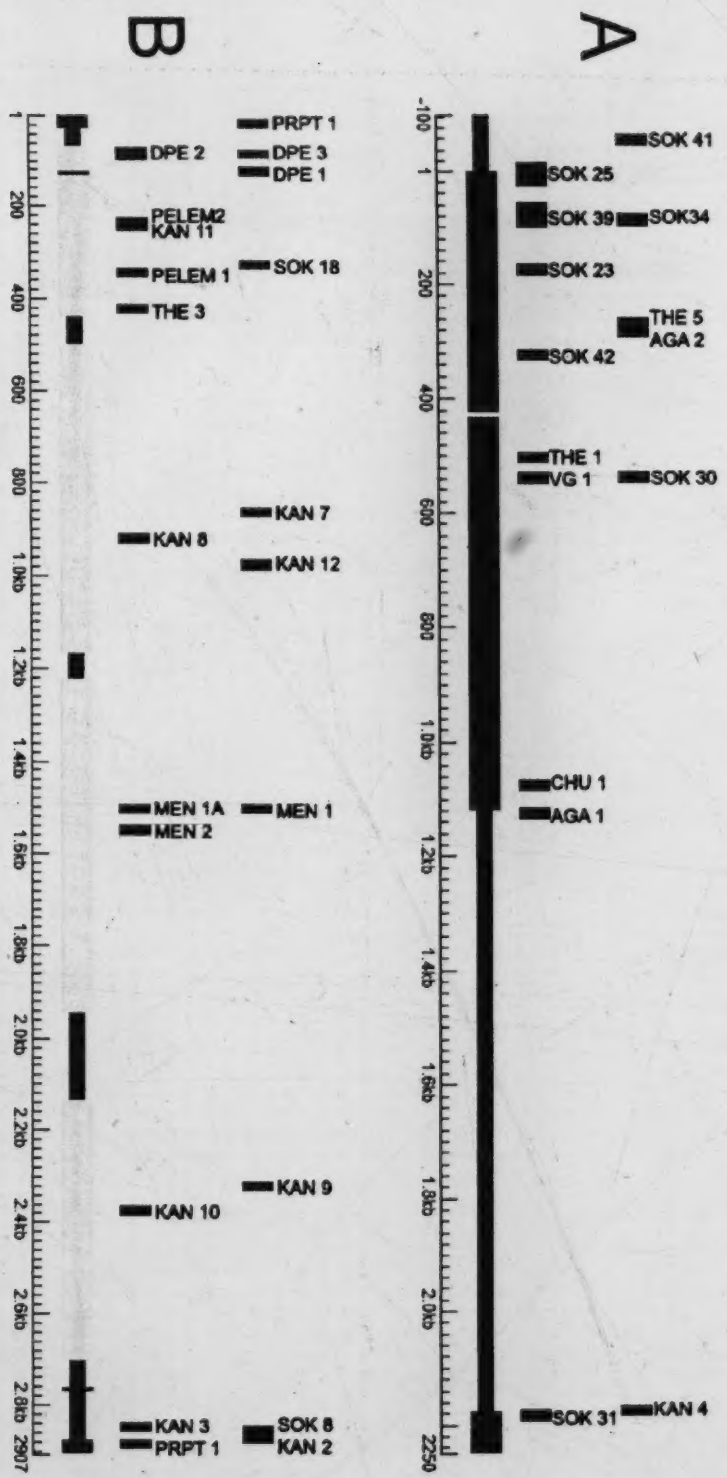

B
PRPT 1
DPE 2
DPE 3
DPE 1
PELEM2
KAN 11
PELEM 1
SOK 18
THE 3
KAN 7
KAN 8
KAN 12
MEN 1A
MEN 2
MEN 1
KAN 9
KAN 10
KAN 3
PRPT 1
SOK 6
KAN 2
1
200
400
600
800
1.0kb
1.2kb
1.4kb
1.6kb
1.8kb
2.0kb
2.2kb
2.4kb
2.6kb
2.8kb
2907
A
-100
1
200
400
600
800
1.0kb
1.2kb
1.4kb
1.6kb
1.8kb
2.0kb
2250
SOK 41
SOK 25
SOK 39
SOK34
SOK 23
THE 5
AGA 2
SOK 42
THE 1
VG 1
SOK 30
CHU 1
AGA 1
SOK 31
KAN 4

Table II-1. Drosophila stocks used in this study.

Stock	Source
w; cn vg$^{21-3}$	John Bell
b cn vg^{21r}	John Bell
y w; vg^{83b27R}*/CyO*	John Bell
b cn vg$^{21-7}$	John Bell
b cn vg^{BG}	John Bell
y^{1} *w*^{1118}	John Locke
Df(2R)vg^{B}*/CyO*	John Locke
w; Bc/Cyo; P[Sal]89D Sb/TM2 Ubx e	John Locke
w; dp; e; P[lacW]ci^{Dplac}	John Locke
y w; e Su(var)3-7^{P49}*/TM2 e*	John Locke
y w; Roi/Sp; Sb/Tb	Shelagh Campbell
b cn vg^{2a33}	Ross Hodgetts
Harwich	Ross Hodgetts
Oregon-R	Ross Hodgetts
y^{1} *w*^{67c23}*; P[EPgy2]AGO2*^{EY04479}	Bloomington Drosophila Stock Center
w**; armi*$^{72.1}$*/TM6C Sb*^{1} *Tb*	Bloomington Drosophila Stock Center
y^{1} *w***; P[lacW]armi*^{1}*/TM3 Sb*^{1} *Ser*^{1}	Bloomington Drosophila Stock Center
ru^{1} *st*^{1} *spn-E*^{1} *e*^{1} *ca*^{1}*/TM3 Sb*^{1} *Ser*^{1}	Bloomington Drosophila Stock Center
b^{1} *pr*^{1} *c*^{1} *px*^{1} *sp*^{1}	Bloomington Drosophila Stock Center
Df(2R)CX1, wg^{12} *b*^{1} *pr*^{1}*/SM1*	Bloomington Drosophila Stock Center
w; cn vg$^{21-3}$*; armi*$^{72.1}$*/TM2 Ubx e*	This Study
w; cn vg$^{21-3}$*; P[lacW]armi*^{1}*/TM2 Ubx e*	This Study
w; cn vg$^{21-3}$*; e Su(var)3-7*^{P49}*/TM2 Ubx e*	This Study
y w P[rep]; cn vg$^{21-3}$	This Study
w; cn vg$^{21-3}$*; Ago2*$^{EY04479w+}$	This Study
y w; +; PHStinger[21-3]/TM2 Ubx e	This Study
cn vg^{21r2}	This Study
cn vg^{21r4}	This Study
cn vg^{21r5}	This Study
cn vg^{21r7}	This Study
cn vg^{21r12}	This Study
cn vg^{21r13}	This Study
cn vg^{21r33}	This Study
cn vg^{21r36}	This Study
cn vg^{21r9}	This Study
y w; cn vg^{21r7}	This Study
y w; cn vg^{21r7}*; Sb/Tb*	This Study
b cn vg^{21}*; P[Sal]89D Sb/TM2 Ubx e*	This Study
w; cn vg$^{21-3}$*; P[Sal]89D Sb/TM2 Ubx e*	This Study
b cn vg^{2a33}*; P[Sal]89D Sb/TM2 Ubx e*	This Study
w; cn vg^{21r7}	This Study
b^{1} *P[21-3]60B4*	This Study
b^{1} *vg*$^{21r7-1r}$ *px*^{1} *sp*^{1}	This Study
vg^{21r7-r} *px*^{1} *sp*^{1}	This Study
w; b cn vg^{BG}*; ru*^{1} *st*^{1} *spn-E*^{1} *e*^{1} *ca*^{1}*/TM2 Ubx e*	This Study
w; cn vg$^{21-3}$*; ru*^{1} *st*^{1} *spn-E*^{1} *e*^{1} *ca*^{1}*/TM2 Ubx e*	This Study
w; cn vg^{21r36}*; P[lacW]ci*^{Dplac}	This Study

Table II-2. Synthesized oligonucleotides used in this study.

Name	Location	Orientation	Sequence (5' to 3')
AGA1	bases 1119 to 1139 of vg	reverse	ttatcgaacccgcacttacca
AGA2	bases 278 to 297 of vg	forward	cttggcaactgcagcacact
CHU1	bases 1069 to 1091 of vg	reverse	agatgttcagtgttgcttgcgta
DPE1	bases 117 to 138 of P	forward	tttgaaaacattaacccttacg
DPE2	bases 73 to 102 of P	reverse	cacacaaccttttcctctcaacaagcaaacg
DPE3	bases 78 to 97 of P	forward	gcttgttgagaggaaaggtt
KAN1	spans P promoter (for S1)	reverse	cctttcctctcaacaagcaaacgtgcactga atttaagtgtgcgagga
KAN2	bases 2868 to 2886 of P	forward	gtctcttgccgacgggacc
KAN3	bases 2841 to 2858 of P	reverse	acgtatgcttgcaataag
KAN4	bases 2171 to 2190 of vg	forward	cttccagttggcaattcttc
KAN7	bases 856 to 876 of P	forward	gatggagttgatgacgccgac
KAN8	bases 912 to 933 of P	reverse	gtcatactcgaaggcagcagcg
KAN9	bases 2314 to 2335 of P	forward	gagttgacggaggatgcgatgg
KAN10	bases 2367 to 2389 of P	reverse	cttttactttgtcactgattctc
KAN11	bases 233 to 256 of P	reverse	cacccaaggctctgctcccacaat
KAN12	bases 969 to 994 of P	forward	tgtccaactggctattgttcgtggtc
KAN13	bases -274 to -251 of P insert at 2L33	forward	cgacgcgcaggaaaaccaaataga
KAN14	bases 452 to 476 of P insert at 2L33	reverse	gccagcggatatgttgcaatttgtg
KAN15	bases -167 to -147 of P insert at 2R50	forward	attccgatgcgagcgcgactg
KAN16	bases 69 to 93 of P insert at 2R50	reverse	ctccaccacttccatttccgctctc
KAN16B	bases 154 to 175 of P insert at 2R50	reverse	cccgttagcagcacccccacac
KAN17	bases -320 to -295 of P insert at 2R60	forward	aagcgagcacggcgtaaaaatcaatg
KAN18	bases 41 to 63 of P insert at 2R60	reverse	cggtcgcagggctgggctattac
MEN1	bases 1496 to 1517 of P	forward	caccgctagctcgatcagacgc
MEN1A	bases 1496 to 1517 of P	reverse	gcgtctgatcgagctagcggtgg
MEN2	bases 1540 to 1562 of P	reverse	cgcagtttcggtggcatttttcaa
PELEM1	bases 337 to 356 of P	reverse	gagtcgcctcctttttaaatg
PELEM2	bases 227 to 246 of P	reverse	tctgctcccacaattttctc
PRPT1	bases 12 to 31 and 2877 to 2896 of P	both	aacataaggtggtcccgtcg
SOK8	bases 2846 to 2867 of P	forward	tgcaagcatacgttaagtggat
SOK18	bases 321 to 340 of P	forward	cctgcaaaaggtcagacatt
SOK23	bases 169 to 188 of vg	reverse	tggagtgcgtctcaattgtg
SOK25	bases -12 to 30 of vg	reverse	gcgccaactgtgtcaacagtgtttgcagtgcc cccgcaatcc
SOK30	bases 534 to 553 of vg	forward	cgatttcagcgtaaacgaga
SOK31	bases 2178 to 2197 of vg	reverse	ttccgacgaagaattgccaa
SOK32	bases 82 to 104 of vg	reverse	ggaattcgcattcatccggggcc
SOK34	bases 82 to 104 of vg	forward	ggccccggatgaatgcgaattcc
SOK39	bases 60 to 83 of vg	reverse	ccttctacttcgtccttttgttcc
SOK41	bases -61 to -42 of vg	forward	acttcgcggttgaggagggc
SOK42	bases 317 to 337 of vg	reverse	cgcacggcagtcattggaccg
THE1	bases 497 to 516 of vg	reverse	atcccgcgcggcggtgagag
THE3	bases 415 to 434 of P	reverse	tgtactcccactggtatagc
THE5	bases 262 to 281 of vg	forward	aatcaagtgggcggtgcttg
VG1	bases 534 to 553 of vg	reverse	tctcgtttacgctgaaatcg

Literature Cited

KOETSIER, P. A., J. SCHORR and W. DOERFLER, 1993 A rapid optimized protocol for downward alkaline Southern blotting of DNA. Biotechniques **15:** 260-262.

RUBIN, G. M., and A. C. SPRADLING, 1982 Genetic transformation of Drosophila with transposable element vectors. Science **218:** 348-353.

WILLIAMS, J. A., J. B. BELL and S. B. CARROLL, 1991 Control of Drosophila wing and haltere development by the nuclear vestigial gene product. Genes Dev **5:** 2481-2495.

Chapter III: Primary and secondary events which revert or repress the phenotype of a *P* element induced *vestigial* allele in *Drosophila melanogaster*.[1]

Introduction

The production of novel mutations is at the very heart of genetic research, and in *Drosophila melanogaster*, one exceedingly useful mutagen is the *P* element. The use of this transposon represents a compromise between the ease with which the mutants can be successfully cloned, and the limitation that transposition preferentially targets some locations in the genome, while avoiding others. Another property of *P* elements making them suitable for mutagenesis is that they preferentially transpose into the 5' end of genes, likely due to some mechanism of targeting transcriptionally active chromatin (KELLEY *et al.* 1987). The basis of *P* element mutagenesis relies on their ability to encode both a transposase (RIO *et al.* 1986), and a repressor (MISRA and RIO 1990; RIO *et al.* 1986; ROBERTSON and ENGELS 1989). Transposase can only be produced in the female germline, as only there does the pre-mRNA become properly spliced (SIEBEL and RIO 1990), while in somatic tissues the third intron remains in the mature mRNA to yield production of a truncated protein (RIO *et al.* 1986). Full *P* elements can therefore only transpose in the germ line, but this can be prevented if the developing embryo contains a sufficient amount of *P* element repressor - acquired through the oocyte - to reduce or prevent transcription from the *P* promoter (LEMAITRE and COEN 1991). With this somewhat simplified molecular explanation, we can understand why reciprocal matings can yield very different results. If the female contains *P* elements, and the male does not, there will be inheritance of repressor, and no new transpositions; conversely, if the

[1] A version of this chapter has been published as Anderson *et al.* 2006. Genome, 49: 1184-1192

female is free of *P* elements, but the male contains at least one full *P* element, then there will be transposase production, and mobilization to new locations in the genome. The phenomenon produced by these transpositions include sterility, pupal lethality, and recombination in males and are collectively called hybrid dysgenesis (KIDWELL *et al.* 1977).

Prediction of the behavior of *P* elements in crosses depends upon the existence of three different types of flies. Stocks or flies which contain no *P* elements are described as having M cytotype, P cytotype designates those which contain *P* elements capable of producing transposase and Q cytotype (sometimes also referred to as M' cytotype) is used to describe flies with *P* elements that will not mobilize when crossed to M cytotype, nor allow the mobilization of elements when mated to P cytotype (ENGELS 1983). Population studies have revealed that while flies in the wild nearly all contain *P* elements, the *P* element has only recently invaded the *D. melanogaster* genome. This is fortuitous as numerous stocks established before 1950 do not contain any *P* elements, allowing researchers to produce dysgenic crosses (KIDWELL 1983). An important question faced by population geneticists is that if *P* element transposition is so mutagenic, would not the accumulation of mutations by the average 30-50 copies of *P* elements per P strain genome produce too many deleterious effects to allow the progression of *P* elements throughout the species? We hope to add a piece of understanding to this puzzle by investigating how a *P* element induced deleterious mutation of the wing determining gene *vestigial* (*vg*) can be repressed by the presence of additional *P* elements.

Vestigial is one of the classic *Drosophila* mutants first described by Morgan (BRIDGES and MORGAN 1919). The molecular description of a *vg* mutant came 75 years

later (WILLIAMS and BELL 1988) with the recovery of a cryptic *P* element-induced mutant designated vg^{21} (WILLIAMS and BELL 1988). This allowed the locus to be cloned and molecularly characterized (WILLIAMS *et al.* 1990; WILLIAMS and BELL 1988). More severe secondary alleles of vg^{21} have been generated in dysgenic crosses, including $vg^{21\text{-}3}$ (WILLIAMS *et al.* 1988) and vg^{2a33} (HODGETTS and O'KEEFE 2001). All *P* element-induced *vg* alleles have their *P* element inserted in the same place in the 5'-untranslated region of the first exon, with the transposon promoter running in parallel with that of the *vestigial* promoter (HODGETTS and O'KEEFE 2001; WILLIAMS *et al.* 1988). An unusual feature of the *P* element alleles of *vg* is that the mutant phenotype of some alleles is suppressed in the presence of *P* elements that encode a repressor whereas other alleles are insensitive to *P* repressor. Sensitivity to *P* repressor, or the lack of it, must be caused by differences in the internal sequences of the *P* elements, since the insertion point and orientation of the *P* element in every allele is the same.

Because the continuous variation of wing area among *vg* alleles is a quantifiable reporter of *vg* function (NAKASHIMA-TANAKA 1968a), we set out to create additional secondary *P* element alleles at *vg* to better understand the relationship between the internal sequence of an element at *vg* and the phenotype which results. In a preliminary control experiment, we attempted to replace a $vg^{21\text{-}3}$ *P* element (P[21-3]) with the vg^{21} *P* element (P[21]) supplied on a plasmid (KEELER *et al.* 1996). A replacement event would result in a reversion of the strong $vg^{21\text{-}3}$ phenotype. The method used microinjection of a plasmid encoding *P* element transposase together with a plasmid containing P[21] into a $vg^{21\text{-}3}$ host embryo to effect germline transformation. Amongst the progeny that exhibited a reversion of the strong $vg^{21\text{-}3}$ phenotype, we expected to find precise excisions

of the vg^{21-3} element, remnants of *P* element ends and internally deleted elements with complete ends resulting from imprecise excisions, and replacements of the P[21-3] element with the P[21] element (GLOOR *et al.* 2000; KEELER *et al.* 1996; STAVELEY *et al.* 1995). While no replacements were observed, we did recover several revertants of the vg^{21-3} phenotype. While most of these fell into the expected classes, several novel events were recovered that promise to shed light on the unusual mechanism of *P* mutagenesis at *vestigial*.

Results

Microinjection to replace vg^{21-3} with vg^{21}

All of the revertant stocks characterized in this study were recovered from an attempted *P* element replacement experiment. Our objective to replace the unmarked *P* element present in vg^{21-3} with that present in vg^{21} necessitated some preliminary characterization of the phenotypes seen in both M and P cytotypes (Figure III-1). The vg^{21} allele is cryptic and homozygotes show no wing defect (Figure III-1A). Slight nicking of the wing is evident in heteroallelic combinations with the null allele vg^{B} (Figure III-1B). Obviously the suppressive effects of the P repressor cannot be scored in vg^{21} homozygotes (Figure III-1C) but they are evident in the wing phenotype of the vg^{21}/vg^{B} genotype (compare Figure III-1D with Figure III-1B). The phenotypes of vg^{21-3}/vg^{21-3} and that of vg^{21-3}/vg^{B} are severe in the absence of any repressor (Figure III-1E and III-1F). Addition of repressor fully restores the wing phenotype when vg^{21-3} is homozygous (Figure III-1G) and almost completely restores the phenotype when heterozygous with a null allele (Figure III-1H).

To sequence the vg^{21} and vg^{21-3} *P* elements they were PCR amplified from genomic fly DNA as single fragments and sequenced using various primer sets located along the *P* element sequence to achieve full and overlapping sequence coverage. The sequence of vg^{21} had been previously published (WILLIAMS *et al.* 1988), but the vg^{21-3} allele had been only described by Southern analysis. Figure III-2 shows the results of this sequencing data in the context of a full wild type *P* element. The P[21] element has a deletion of bases 382 to 2603 but is otherwise identical in sequence to a full *P* element. The P[21-3] element has been found to contain a deletion of bases 1862 to 2015 with a 2 base AT insertion at the deletion site and of particular note is a single A to T transversion in P[21-3] at base 32. This polymorphism allows the determination of the origin of a *P* element from a P[21] element (A at base 32) or P[21-3] (T at base 32) regardless of the extent of any internal deletions.

Once this preliminary groundwork had been laid, we commenced with microinjection of pGEM-T-vg^{21} and pUChsΔ2–3 plasmids into vg^{21-3} hosts. To replace one *P* element with another at *vestigial* we relied on previous work showing that homologous recombination is possible in *Drosophila melanogaster*. By mobilizing a *P* element to create a double strand break at its genomic site, Keeler *et al* (1996) showed that the break could be repaired from a plasmid template. The injection of the repair template and transposase into the germline of *w* ; vg^{21-3} flies created germline mosaics containing alterations of the *vestigial* locus which were then crossed to a vg^{B}/*CyO* stock to test for wing phenotype over the deficiency chromosome vg^{B}. Possible replacements were then selected for by only taking flies which had wild type or nicked wings in contrast to the strong mutant wings expected if no event had taken place at *vg*. Once

these flies were bred to homozygosity we did several tests to determine the type of *P* element at *vg*.

Simple revertant alleles

Nine unique alleles were created using our replacement technique, and six of those alleles proved easily amenable to cloning. A PCR reaction spanning the *P* element insertion site of all the *vg* alleles was used to clone and then sequence the six alleles with elements less than 1 kbp in length. Figure III-3 shows these new alleles in relation to the P[21] element. The four alleles vg^{21r5}, vg^{21r12}, vg^{21r13} and vg^{21r33} may be easily explained as excisions of the *P* element which have left behind "remnants" of sequence, usually approximating half of each 31bp repeat (STAVELEY *et al.* 1995). In addition to the remnants of 31 bp repeats, vg^{21r13} contains 28 bases of sequence (shown in green in Figure III-2) which sequence analysis revealed as likely arising by copying in a region of 2L:32C1 to repair the double strand break using non homologous sequence. The vg^{21r2} allele appears to be an incomplete excision of the P[21-3] from *vg*, which leaves behind 16 bases of the 5' 31 bp repeat, then an insertion of TAA and then bases 2530 to 2907 of a full element at the 3' end. The bases shown in cyan of vg^{21r2} are present in P[21-3] but not P[21] which tells us the origin of the element is from a deletion of most of P[21-3]. A more complex configuration was found when sequencing vg^{21r36}. A preliminary look at the PCR product of this allele shows a *P* element approximately the same size as the one in vg^{21} but sequencing and alignment showed that an additional 48 bases were present past the 5' end of the deletion in the P[21] element (bases marked in cyan in Figure III-3) and the deletion went 77 bases further than the P[21] element. The P[21r36] element also

contained the polymorphism at base 32 strongly suggesting the element arose from the P[21-3] element.

Complex revertant alleles

Three of the nine alleles selected for by reversion of wing phenotype proved to be more difficult to analyze. One allele, vg^{21r9}, was discarded early in our experiments as it proved to be homozygous lethal and likely had a secondary insertion into an essential gene. The two remaining alleles were designated vg^{21r4} and vg^{21r7}. Both alleles proved initially difficult to PCR amplify from genomic DNA as our protocol was optimized for inserts of less than 1 kbp. However a more thorough study yielded insights into the composition of the *vg* locus at both alleles. While long PCR allowed us to clone and sequence vg^{21r7} as a single 3 kbp genomic fragment, no optimizations would yield a complete insert of the vg^{21r4} stock. We reasoned that by using a combination of both *vg* and *P* element primers we could determine why the insert in this stock was so intractable to a single PCR. Sequencing a PCR product of SOK34 to KAN3 showed the first 648 bases of the 5' end of the *P* element is identical to P[21], including the polymorphism at base 32 and deletion from *P* element base 382 to 2602. When trying to clone from the 3' end however, we find that the sequence is that of P[21-3] for at least 1898 bases which is all sequence 3' to KAN7. The apparent discrepancy in size from each end can be explained by a P[21] element inserting next to the 5' end of vg^{21-3}. While vg^{21r4} gave us very interesting results in that we were not expecting tandem insertions, vg^{21r7} is interesting for a completely different reason. Sequencing of the entire element from a

single genomic fragment showed no difference from the P[21-3] element that was present at *vg* in our injection stock.

Phenotypic scoring of new alleles

As a P[21-3] element at *vg* produces a strong phenotype in the absence of P repressor but a wild type phenotype in the presence of P repressor we reasoned that there may be additional *P* elements in the new vg^{21r7}. If this was the case, putting the new stock over a *vg* null allele would not only decrease the number of potential repressor producing *P* elements in half, but also provide only a single, *P* inserted copy of *vg* with which to produce VG product. Each of the new alleles was crossed to vg^{B}/CyO and approximately 50 wings or more were scored for wing phenotype. The phenotypic scale we used has been modified from Nakashima-Tanaka (NAKASHIMA-TANAKA 1968b) such that wings with no nicking are scored as wild type rather than given a value of 1-5 for size. The remaining values of 6-31 were scored according to the original scale. Figure III-4 shows the results of the new stocks with the corresponding wing phenotype illustrated underneath. As expected for most of our homozygous wild type stocks, they showed mostly wild type wings with some occasional minor nicking. The results from vg^{21r7}/vg^{B} showed a Poisson distribution from wild type to mild and moderate nicking with over 90% of the wings scored having a non-wild type phenotype. Unexpectedly, vg^{21r36}/vg^{B} showed a strong phenotype reminiscent of that of vg^{21-3} as no wild type wings were seen and all wings were considered moderate to strong in phenotype.

Southern analysis

With the results from vg^{21r36} and vg^{21r7} showing interesting cryptic phenotypes, we chose to do Southern analysis on these stocks to help determine if additional *P* elements were present which could act as a source of repressor. Control stocks which we used were vg^{21-3} and vg^{21} as they were the initial stocks of interest and vg^{B}/CyO to serve as an M cytotype stock. Figure III-5 shows the result of 5μg of EcoRI digested genomic DNA samples probed with a randomly labeled PCR fragment of bases 117 to 933 of a full *P* element. The control lanes are vg^{B}/CyO which shows no *P* elements and the original vg^{21} and vg^{21-3} stocks used in this experiment. All of our *P* element carrying stocks show a high molecular weight doublet of bands which by intensity corresponds to less than 260 bp of sequence complementary to our probe. In the vg^{21r7} lane we are able to see that this stock still contains the appropriate band of 2044 bp for a P[21-3] sized element at *vg* and an additional 3 strong bands of approximately 1.9 kbp, 2.8 kbp and 3.6 kbp in size which are new *P* element insertions. The only difference between vg^{21-3} and the new allele vg^{21r36} is the loss of a 2044 bp band and acquisition of a 1356 bp band as expected if the P[21r36] sequence was inserted into *vg* (Figure III-4).

Discussion

The initial purpose of this series of experiments was to test the experimental methods of homologous *P* element replacement by plasmid injection. Although no exact replacement of P[21-3] with P[21] at *vg* was seen, we did recover at least one case of targeted secondary insertion into the same genomic location in the allele vg^{21r4}. We recovered and sequenced 648 bases of the 685 bp P[21] element inserted into the same

genomic location of vg; however, the last 37 bases could not be used to amplify a fragment to sequence as 31 bases of that sequence are at least duplicated and likely quadruplicated as the *P* element terminal inverted repeats. By attempting to sequence from the 3' end of the insertion site into the *P* element we found that at least 1898 bp of the P[21-3] sequence was also present. Further sequencing upstream was once again hindered by the presence of the same primer sites in P[21-3] as in the 5' P[21] as shown in Figure III-6, where we show a potential origin of this complex allele. In order for transposase to cut the 5' end of the *P* element at *vg* it must first cut the 3' end of the *P* element (BEALL and RIO 1997) (Figure III-6A). As shown in Figure III-6B the 2 double strand breaks produced by excision of the P[21-3] element from *vg* are then repaired using separate templates of pGEM-T-vg^{21} at the 5' end and a homologous chromosome or sister chromatid containing P[21-3] is used to repair the 3' double strand break. The 2 halves of the chromosome must undergo crossing over to allow for the recovery of an intact Chromosome II in the vg^{21r4} stock as depicted in Figure III-6C. The black bars in Figure III-6D indicate the extent of sequenced *P* element as compared to the unique *P* element sequence shown with a red bar. The remainder of *P* element sequence not underlined with a red bar is present in both of the tandem elements. The duplication of *P* element sequences common to both P[21] and P[21-3] made it difficult to clone and sequence the region, but sequencing from the unique surrounding genomic region allowed the verification of P[21] sequence at *vg*.

The double insertion of *P* elements at *vg* is similar to two alleles of *singed*, sn^{weak} and $sn^{very\ weak}$, in which two elements that are tandemly inserted allow for a weak phenotype when both are present (NITASAKA and YAMAZAKI 1988; ORTORI *et al.* 1994).

When transposase is introduced into a sn^{weak} or $sn^{very\ weak}$ stock, the loss of one of the elements in each pair allows for the remaining element to exert a strong effect on *singed* gene function. It is possible that if the P[21] element can be removed by transposase, leaving only P[21-3] at *vg*, the acquisition of a strong wing phenotype could be used as a test for the effectiveness of transposase source.

Although we did not achieve replacement of one *P* element with another, we did manage to target a *P* element contained on a plasmid to a genomic location using a helper plasmid which encoded transposase. This targeted integration process can be used to better understand the mechanism of gene impairment at *vg* by *P* elements as we can modify sequences in the transposon we believe are important for determining the strength of wing phenotype and response to repressor and ascertain those effects by reintroducing those elements at *vg*. This method may also prove useful to other researchers who wish to insert a specific *P* element, perhaps containing specific markers into a location of interest for which there is already a *P* element insertion. Due to our selection of a weak vestigial phenotype we biased the experiment to recover excision events, but if a replacement experiment was carried out in which a strong or unique phenotype was selected for, then the rate of false positives should be greatly minimized.

Although these experiments gave some unexpected results, the production of novel mutants is beneficial to our understanding of the *P* element and *vg* interaction. Unlike any *P* element allele of *vg* known, vg^{21r36} is a weak mutant when homozygous but a strong mutant in combination with a deficiency. The further investigation of how this phenotype manifests when fewer than 100 bases are different from that of vg^{21} may help us understand a critical region important in determining wing development in this system.

It is mutants like these which highlight regions of potentially important *P* element sequence that will allow us to better design *P* elements for replacement to test our understanding of the intricacies of *P* and *vg* interaction.

The acquisition of repressor making elements in the vg^{21r7} stock was unexpected as there was no source for repressor present in the experiment. Looking at the three *P* elements initially present in our transgenic flies, the P[21] and P[Δ2-3] elements on plasmids and the P[21-3] element at *vg*, we can try to identify where repressor could come from. The P[21] element is a highly truncated element, with the deletion spanning from exon 0 to exon 3, whereas KP, D50 and all other known Type II repressors contain all of exon 0 and approximately half of the larger exon 1 before any sequences are deleted (ENGELS 1996). Exon 1 has been shown to contain a leucine zipper motif which is important for repressor activity and is lacking in the P[21] element (ANDREWS and GLOOR 1995). We therefore doubt that the P[21] element has any potential to make repressor. While the P[Δ2-3] element could be internally deleted to give either a Type I or Type II repressor, it is very unlikely to do so since it cannot transpose. The lack of a 3' 31 bp repeat renders it immobile as the 3' end of the element is the initial site of action of transposase (BEALL and RIO 1997). This leaves us the P[21-3] element as a potential source of repressor making elements. Although the P[21-3] element does not contain enough sequence to fall into the classification of a Type I element, it could undergo further internal deletion to give a perfect KP or other Type II repressor element. It is possible that one or all of the additional elements seen in the vg^{21r7} stock has undergone internal deletion to allow the production of repressor. Further investigation into the location and composition of these unmarked *P* elements should verify their origin from

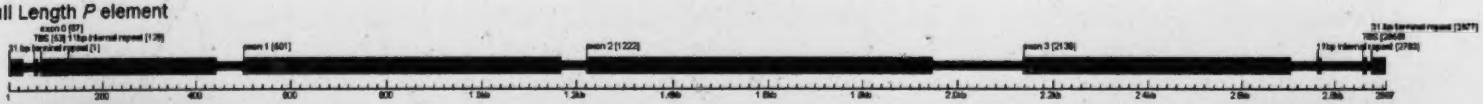
the P[21-3] element. The presence of multiple Type II elements on third chromosomes designated KP-U and KP-D has been shown to repress the vg^{21-3} phenotype (Dan Bushey, personal communication) making the possibility of P[21-3] producing a Type II which could repress the phenotype the most likely scenario.

It is easy to see how a *P* element can cause gene mutation in cases where the insertion interrupts the coding region of the gene, but in cases where the effect of the insertion is reliant on a *P* encoded repressor we are left with a very incomplete understanding. In the specific case of vg^{21-3}, where a severe phenotype is rendered wild type by the repressor, we hope that the targeting of modified *P* elements may lead us to understand what sequences are important for both severity of the phenotype and responsiveness to repressor. It is with this investigation of the interactions of *P* elements at *vg* that we hope to better appreciate the subtle ways in which a transposable element can modify gene expression. By understanding how a *P* element can both produce a phenotype with one insertion and then repress that phenotype with a secondary insertion we may better understand how they have achieved such success in invading the *Drosophila melanogaster* genome.

Figure III-1: Phenotypic response of vg^{21} and vg^{21-3}. P repressor was produced by the P[Sal]89D element on chromosome III. Wings of vg^{21}/vg^{21} **(A)** and vg^{21}/vg^{B} **(B)** in the absence of P repressor. Wings of vg^{21}/vg^{21} **(C)** and vg^{21}/vg^{B} **(D)** in the presence of P repressor. Wings of vg^{21-3}/vg^{21-3} **(E)** and vg^{21-3}/vg^{B} **(F)** in the absence of P repressor. Wings of vg^{21-3}/vg^{21-3} **(G)** and vg^{21-3}/vg^{B} **(H)** in the presence of P repressor. All wings are shown to the same scale.

Figure III-2. Sequence of initial *P* elements. The sequence of a full length *P* element is represented with its four exons in green, 31 bp repeats shown in pink, 11 bp repeats shown in orange and transposase binding sites shown in red. The deleted regions of the 3 preexisting alleles used in this study are shown below the full element. vg^{21} has a deletion of bases 380 to 2602. vg^{2a33} (not shown) has the same deletion of 380 to 2602 as vg^{21} and an additional deletion of bases 2755 to 2800 with 4 Cs inserted in the 3' deletion. vg^{21-3} has a deletion of bases 1864 to 2015 and also contains a A to T transversion at base 32.

Full Length *P* element



vg^{21} *P* element

vg^{21-3} *P* element

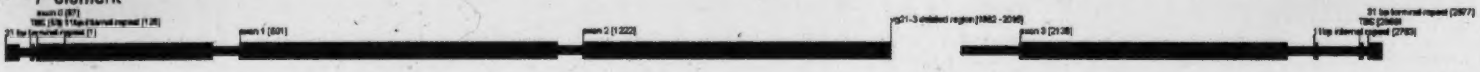


Figure III-3: Alignment of six new alleles with P[21]. An alignment of the sequences of vg^{21r2}, vg^{21r5}, vg^{21r12}, vg^{21r13}, vg^{21r33} and vg^{21r36} with vg^{21} is shown with *P* element sequence present in the new stocks but not in P[21] shown in cyan and novel sequence shown in green. TBS is the transposase binding site and Pp is the *P* element promoter.

Vg21
Vg21r2
Vg21r5
Vg21r12
Vg21r13
Vg21r33
Vg21r36
P insertion site
31bp terminal repeat
TSS
Ep
118p internal repeat
vg21 deletion
P terminator
118p internal repeat
TSS
31bp terminal repeat
P insertion site

Figure III-4. New alleles in heterozygous combination with a null allele. Each of the new *vg* stocks vg^{21r2}, vg^{21r5}, vg^{21r4}, vg^{21r12}, vg^{21r13}, vg^{21r33}, vg^{21r7} and vg^{21r36} was scored in combination with vg^{B} at 25°C. The wing phenotypes scored correspond to the illustrated wing scale shown below the chart and in Figure I-3.

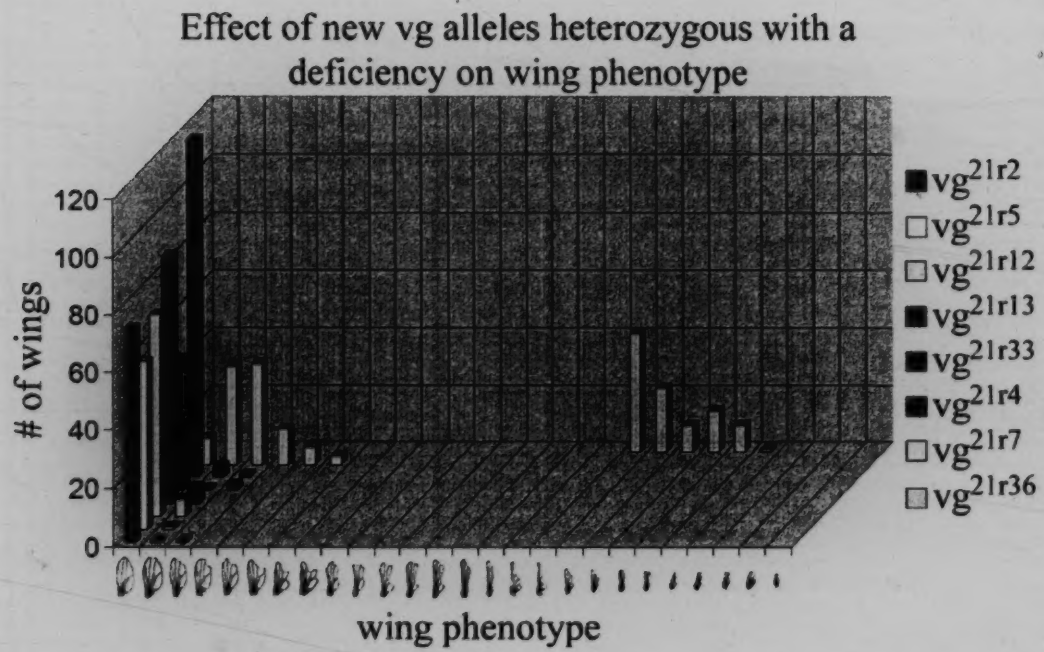

Effect of new vg alleles heterozygous with a deficiency on wing phenotype
# of wings
120
100
80
60
40
20
0
wing phenotype
vg21r2
vg21r5
vg21r12
vg21r13
vg21r33
vg21r4
vg21r7
vg21r36

Figure III-5: Southern hybridization of novel stocks created in the *P* element replacement experiment. Genomic DNA for each of the indicated stocks was digested with *Eco*RI, run on a 1% agarose gel and transferred using a downward alkaline system (KOETSIER *et al.* 1993) to a positive nylon membrane and probed using a randomly labeled PCR product corresponding to bases 117 to 933 of a full *P* element. Sizes of fragments were calculated by comparing to size standards visualized with ethidium bromide on the initial agarose gel.

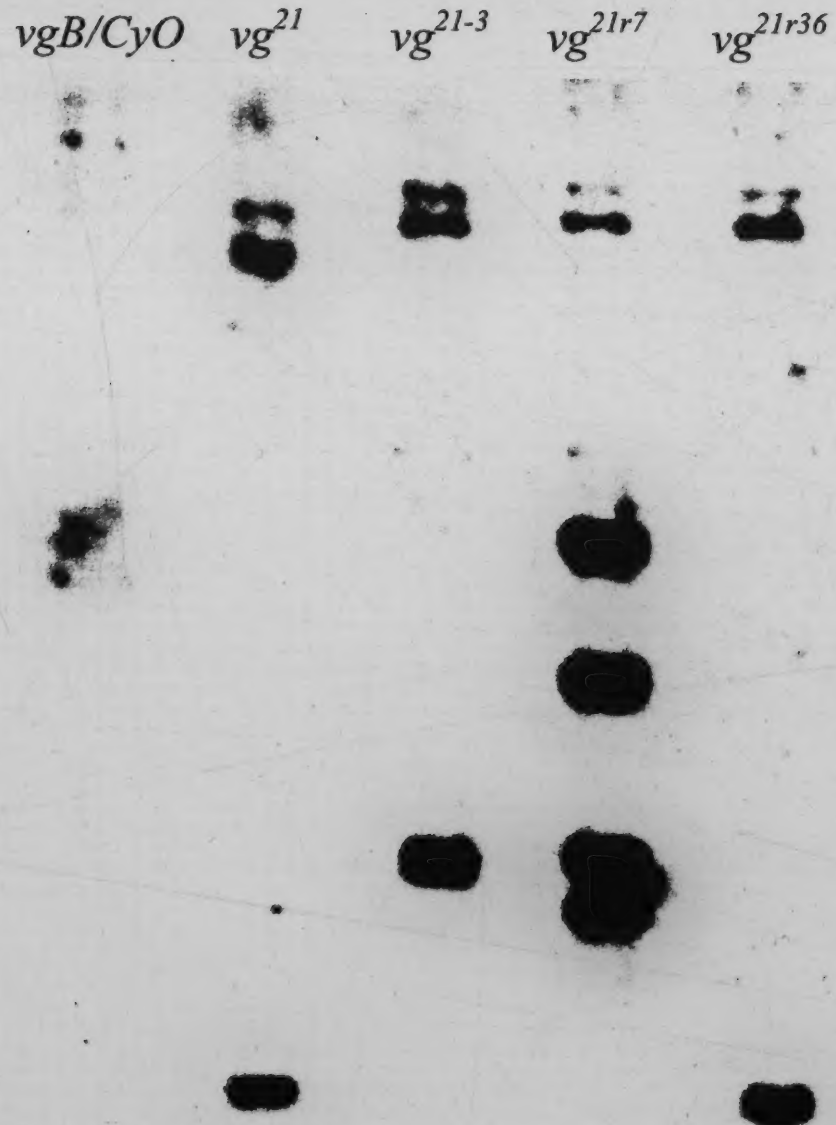
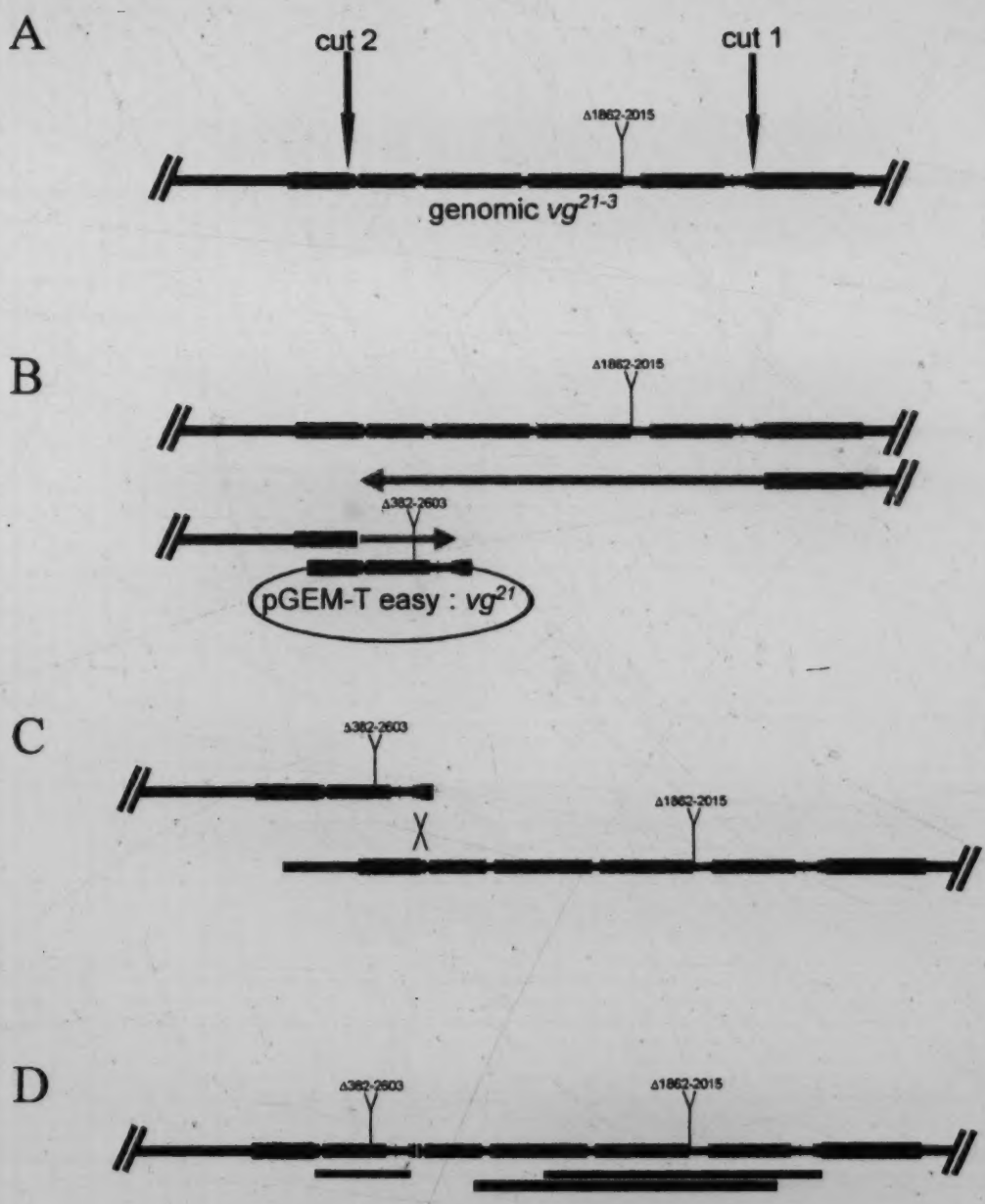

vgB/CyO
vg21
vg21-3
vg21r7
vg21r36

Figure III-6. Model of targeting P[21] into vg^{21-3} to create vg^{21r4}. **A.** Transposase creates a double strand break. **B.** Repair of the double strand break using 2 homologous templates. **C.** Crossover of the repair products. **D.** Schematic of the sequenced vg^{21r4} element. The verified sequence is shown by the black bar while sequence which is unique to only one element is shown with the red bar.


A
cut 2
cut 1
Δ1862-2015
genomic vg21-3
B
Δ1862-2015
Δ382-2603
pGEM-T easy : vg21
C
Δ382-2603
Δ1862-2015
D
Δ382-2603
Δ1862-2015

Literature Cited

ANDREWS, J. D., and G. B. GLOOR, 1995 A role for the KP leucine zipper in regulating P element transposition in Drosophila melanogaster. Genetics **141:** 587-594.

BEALL, E. L., and D. C. RIO, 1997 Drosophila P-element transposase is a novel site-specific endonuclease. Genes Dev **11:** 2137-2151.

BRIDGES, C. B., and T. H. MORGAN, 1919 Contributions to the genetics of Drosophila melanogaster. Publs Carnegie Instn **278:** 123-304.

ENGELS, W. R., 1983 The P family of transposable elements in Drosophila. Annu Rev Genet **17:** 315-344.

ENGELS, W. R., 1996 P Elements in Drosophila, pp.

GLOOR, G. B., J. MORETTI, J. MOUYAL and K. J. KEELER, 2000 Distinct P-element excision products in somatic and germline cells of Drosophila melanogaster. Genetics **155:** 1821-1830.

HODGETTS, R. B., and S. L. O'KEEFE, 2001 The mutant phenotype associated with P-element alleles of the vestigial locus in Drosophila melanogaster may be caused by a readthrough transcript initiated at the P-element promoter. Genetics **157:** 1665-1672.

KARESS, R. E., and G. M. RUBIN, 1984 Analysis of P transposable element functions in Drosophila. Cell **38:** 135-146.

KEELER, K. J., T. DRAY, J. E. PENNEY and G. B. GLOOR, 1996 Gene targeting of a plasmid-borne sequence to a double-strand DNA break in Drosophila melanogaster. Mol Cell Biol **16:** 522-528.

KELLEY, M. R., S. KIDD, R. L. BERG and M. W. YOUNG, 1987 Restriction of P-element insertions at the Notch locus of Drosophila melanogaster. Mol Cell Biol **7:** 1545-1548.

KIDWELL, M. G., 1983 Evolution of hybrid dysgenesis determinants in Drosophila melanogaster. Proc Natl Acad Sci U S A **80:** 1655-1659.

KIDWELL, M. G., J. F. KIDWELL and J. A. SVED, 1977 Hybrid dysgenesis in Drosophila melanogaster: a syndrome of aberrant traits including mutation, sterility and male recombination. Genetics **86:** 813-833.

KOETSIER, P. A., J. SCHORR and W. DOERFLER, 1993 A rapid optimized protocol for downward alkaline Southern blotting of DNA. Biotechniques **15:** 260-262.

LEMAITRE, B., and D. COEN, 1991 P regulatory products repress in vivo the P promoter activity in P-lacZ fusion genes. Proc Natl Acad Sci U S A **88:** 4419-4423.

MISRA, S., and D. C. RIO, 1990 Cytotype control of Drosophila P element transposition: the 66 kd protein is a repressor of transposase activity. Cell **62:** 269-284.

NAKASHIMA-TANAKA, E., 1968a The effect of temperature and genetic background on the phenotypic expression of several vestigial strains of Drosophila melanogaster. Genetica **38:** 447-458.

NAKASHIMA-TANAKA, E., 1968b The effects of chemicals on the phenotypic expression of a vestigial mutant in Drosophila melanogaster. Genetica **38:** 459-470.

NITASAKA, E., and T. YAMAZAKI, 1988 Excision of one of two defective P elements as the cause of alternate mutational events (sn+ and sne) of the singed bristle allele snw in Drosophila melanogaster. Jpn J Genet **63:** 303-312.

ORTORI, C. A., D. CHAMBERS and J. F. BROOKFIELD, 1994 The molecular basis of instability of the singed(very weak) mutation in Drosophila melanogaster. Genet Res **63:** 19-26.

RIO, D. C., F. A. LASKI and G. M. RUBIN, 1986 Identification and immunochemical analysis of biologically active Drosophila P element transposase. Cell **44:** 21-32.

ROBERTSON, H. M., and W. R. ENGELS, 1989 Modified P elements that mimic the P cytotype in Drosophila melanogaster. Genetics **123:** 815-824.

SIEBEL, C. W., and D. C. RIO, 1990 Regulated splicing of the Drosophila P transposable element third intron in vitro: somatic repression. Science **248:** 1200-1208.

STAVELEY, B. E., T. R. HESLIP, R. B. HODGETTS and J. B. BELL, 1995 Protected P-element termini suggest a role for inverted-repeat-binding protein in transposase-induced gap repair in Drosophila melanogaster. Genetics **139:** 1321-1329.

WILLIAMS, J. A., A. L. ATKIN and J. B. BELL, 1990 The functional organization of the vestigial locus in Drosophila melanogaster. Mol Gen Genet **221:** 8-16.

WILLIAMS, J. A., and J. B. BELL, 1988 Molecular organization of the vestigial region in Drosophila melanogaster. Embo J **7:** 1355-1363.

WILLIAMS, J. A., S. S. PAPPU and J. B. BELL, 1988 Molecular analysis of hybrid dysgenesis-induced derivatives of a P-element allele at the vg locus. Mol Cell Biol **8:** 1489-1497.

Chapter IV: Non-canonical *P* element repressor-making elements in *Drosophila melanogaster.*

Introduction

The ability of some *P* elements to repress the effect of other *P* elements is very robust, as evidenced by the wide range of repressor making elements observed in *Drosophila melanogaster* (BLACK *et al.* 1987; GLOOR *et al.* 1993; ROBERTSON and ENGELS 1989; RONSSERAY *et al.* 2001). Repressors which act in the germline will prevent *P* element mobilization and the deleterious effects collectively termed hybrid dysgenesis. In addition to the germline repressor effects, *P* element expression in somatic cells can also be repressed. There are three commonly used assays to detect somatic *P* repression. The first assay, $singed^{weak}$ hypermutability, involves an allele of *singed* which harbors a tandem *P* element insertion (NITASAKA and YAMAZAKI 1988). In the presence of somatic transposase, the weak *singed* phenotype may become wild type or severe depending on which of the two elements mobilizes. If repressor is also present there will be a measurable decrease in the frequency of $singed^{weak}$ allele modifications. The main drawback of this system is the requirement for somatic transposase which can only be expressed using a modified element in which the third intron of the transposase gene has been deleted. A second and also extensively used assay involves a *P* element allele of *vestigial* called $vg^{21\text{-}3}$, that produces a strong vestigial phenotype in which the wings are reduced to stumps (WILLIAMS *et al.* 1988). If repressor-making elements are added to this system, the wing phenotype may be completely restored. The third common assay utilizes modified *P* elements containing a mini-*white* gene; some copies of this element that confer wild type eye color have been demonstrated to produce lighter

colored or variegated eyes with the addition of repressor (BUSHEY and LOCKE 2004). While some repressors prove to be effective for all tests and at several genomic locations, many repressors function only when at specific genomic sites, or in a subset of the repression tests.

Many full length and internally deleted *P* elements have been shown to produce repressor and the various elements have been placed into two groups based on sequence and activity (GLOOR *et al.* 1993). The Type I elements are elements that are nearly full length and systematic deletion from the 3' end has shown that a large decrease in repressor activity occurs if the deletion extends 5' to base 1956. Further deletions would remove the third exon-intron boundary of the *P* element that may be required for the maternal repression effect seen in full length elements due to competition for germline splicing of the pre-mRNA. The Type II elements have been less extensively studied for sequence requirements but are found in most wild *Drosophila* populations (BLACK *et al.* 1987). This second class commonly has over 60% of the *P* element sequence deleted with the endpoints of the deletion located in the second and fourth exons. Two members of this class, the KP element and D50 element, are geographically widespread although other elements with similar deletion endpoints do exist and confer similar activity (ENGELS 1996; RASMUSSON *et al.* 1993).

In this study we describe *P* elements which confer somatic repression but are not readily classified as Type I or Type II repressors. Stocks containing these elements are the result of attempted targeted transposition to replace a *P* element at *vestigial* with a different *P* element. The repressor-making *P* elements which were recovered from these experiments are either too short to be classified as Type II elements as in the case of

P[21r36] or fall into the unclassified area between Type I and Type II elements as seen in P[21-3]. From these results it can be argued that the designation of an element as a Type I or Type II repressor does not adequately reflect the spectrum of repressor elements produced by spontaneous internal *P* element deletion.

Results

vg^{21r7} analysis

One of the lines produced in the mobilization experiment (ANDERSON *et al.* 2006) was designated vg^{21r7} and preliminary work had demonstrated the presence of *P* elements not only at *vestigial* but also at multiple additional genomic locations. The stock produces flies with wild type wings despite the *P* element at *vestigial* being the highly deleterious P[21-3]. Crosses to the *vestigial* deletion chromosome vg^{B} had shown that defective wings could be produced by this line which hinted at the production of repressor by the additional elements acting on the vg^{21-3} allele. In order to study the non-*vestigial P* elements on chromosome II and determine if they were producing repressor their locations had to first be determined.

P element localization

Modified *P* elements containing marker genes such as P[lacZ w^{+}] are commonly localized using an inverse PCR procedure and although the PCR primer sites of the published technique could not be used, it served as a source for altering the primer sites and conditions to allow amplification of a wild type *P* element (Figure IV-1) (REHM 2003). Genomic DNA was digested with *Sau*3A I, ligated under circularizing conditions

and then subjected to PCR of the *P* element and surrounding genomic region. The inverse PCR from KAN11 to KAN12 primers would amplify the genomic sequence 5' to each element to the nearest Sau3AI site (Figure IV-1A). Five bands were produced in an optimized inverse PCR and the purified DNA fragments were sequenced directly (Figure IV-1D). As the PELEM2 sequencing primer lay within 250 bp of the 5' end of the element, very little sequence was required to give the genomic sequence 5' to the element. The sequence was then used for a BLAST search against the *Drosophila melanogaster* genome to find the location of each new insertion. Sequencing of the bands suggested three new insertion sites on chromosome II. The first element was located at 2L-12507808, which corresponds to cytological position 2L33E7, within the largest intron of the 84 kbp gene *bunched*. Another insertion was found at position 2R-9055693, which corresponds to 2R50B1 and gene CG6145. The third insertion occurred at base position 2R-19529683, located at 2R60B4 between genes CG3907 and CG11388. The position of each element relative to the original insertion at *vg* is shown in Figure IV-2B. To confirm the location of each element, the surrounding genomic sequence was used to design PCR primers to span the new genomic sites. Each element was amplified in 2 slightly overlapping halves and those fragments were cloned and sequenced as per Chapter 2 (Figure IV-3). Upon assembly of multiple sequencing reactions the sequence of each element was found to match the sequence of the progenitor P[21-3]. Due to their origin from the vg^{21-3} *P* element these new insertions are named P[21-3]33E7, P[21-3]50B1 and P[21-3]60B4.

Recombinants

The question of whether one of the 3 new inserts was responsible for the repression of the vg^{21-3} wing phenotype or if they all contributed to the repression could not be answered with the current lines. This led to an attempt to recover a chromosome II from the vg^{21r7} line by removing one or more of the non-*vg* P[21-3] elements while retaining the original insert at *vestigial*. Assaying the wing phenotype of the recombinants would allow the contribution of the missing element(s) to wing phenotype repression to be determined. A stock with the markers, $b^1\ pr^1\ c^1\ px^1\ sp^1$, was chosen for recombination studies as the genes *black*, *curved*, *plexus* and *speck* serve as close markers to the 4 elements as seen in Figure IV-2A. After extensive selection for recombinants by back-crossing to the $b^1\ pr^1\ c^1\ px^1\ sp^1$ stock, homozygous stocks of three vg^{21r7} derivatives were produced. In one, the P[21-3]60B4 element had been removed as shown by the non-$curved^1$ wing progeny carrying $plexus^1$ and $speck^1$. As this vg^{21r7}-derived stock is now missing the Far Right *P* element, it will be named $vg^{21r7\text{-}FR}$ (Figure IV-2C). The second stock had both P[21-3]33E7 and P[21-3]60B4 removed as the non-*curved* progeny carried $black^1$, $plexus^1$, and $speck^1$. The absence of the Left and Far Right elements cause this stock to be designated $vg^{21r7\text{-}L\text{-}FR}$ (Figure IV-2D). Due to the proximity of P[21-3]50B1 to *vg*, and the lack of visible markers to select for recombinants, a clean vg^{21-3}-like chromosome was never recovered from the vg^{21r7} stock despite multiple attempts. In addition to the vg^{21-3}–carrying derivatives, the P[21-3]60B4 element was also established on a chromosome devoid of any other *P* elements as it carried $black^1$, $purple^1$, and $curved^1$ (Figure IV-2E). Once proper stocks had been established, a confirmation of the *P* element content of each stock was tested by Southern

analysis. Genomic DNA was extracted from each stock and digested with *Eco*RI which will cut all P[21-3] elements at base 1711. The digested DNA was then run and hybridized according to protocols described in Chapter II and hybridized with a probe specific to bases 117 to 933 of a full-length *P* element. The Southern blot shows that 4 bands of the predicted sizes were seen in DNA from the vg^{21r7} stock (Figure IV-4, Lane A). The new recombinant stocks $vg^{21r7\text{-}FR}$, $vg^{21r7\text{-}L\text{-}FR}$ and P[21-3]60B4 had bands which correspond to the predicted sizes for *P* element locations shown in Figure IV-2 (Figure IV-4, Lanes B – D). Analysis of wings from the 2 stocks retaining the P[21-3] insert at *vg* showed wild type wing morphology. To determine if the loss of *P* elements in each stock altered the wing phenotype in a more sensitive background, each stock was placed over the null allele vg^{B}. Recombinant chromosome carrying females were crossed with with vg^{B}/CyO males and the wings were scored in the progeny on the modified 31 point scale (ANDERSON *et al.* 2006). While the original vg^{21r7} stock yields wings ranging from w.t. to 10 with a peak at 7 on the 31 point scale, the $vg^{21r7\text{-}FR}$ ranges from w.t. to 21 with a peak at 16 and the $vg^{21r7\text{-}L\text{-}FR}$ ranges from w.t. to 10 with most wings being normal (Figure IV-5). Kruskal-Wallis non-parametric ANOVA testing was used to determine that both $vg^{21r7\text{-}FR}$ and $vg^{21r7\text{-}L\text{-}FR}$ did show a statistically significant difference in wing malformation for each population compared to the original vg^{21r7} and also when comparing the two recombinants to each other.

Additional somatic repression

Results of tests described in Chapter III using vg^{21r7} and vg^{21r36} over null alleles showed that these stocks, which produce wild type wings as homozygotes, could produce

a substantial number of flies with mutant wings. This suggests that in the homozygous situation, more *P* elements will lead to repression of the effects of the element at *vestigial*. This repression could exert itself on *P* elements at additional locations, and to test this, a *P* element-borne *white*$^{+}$ transgene was used. A stock was acquired that contains a P[*lacZ white*$^{+}$] insertion on the fourth chromosome near the *cubitus interruptus* gene. Flies of the *w; dp; e;* P[*lacZ w*$^{+}$]*ci* stock have red eyes in the absence of P repressor. Introducing a P repressor into the stock causes a loss of expression of the *P* element borne *white*$^{+}$ gene and the eyes lose pigmentation (BUSHEY and LOCKE 2004).

In the first test of repressor activity, I introduced a single copy of the marked fourth chromosome into a homozygous vg^{21r7} background. As the cross would yield progeny with heterozygous P[*lacZ white*$^{+}$]*ci*/+ and homozygous +/+ fourth chromosomes it was imperative to test for β-galactosidase expression to verify the presence of the P[*lacZ white*$^{+}$]*ci* using a test which did not rely on eye color phenotype which could be repressed by additional *P* elements. The adults were scored for eye phenotype, and also for β-galactosidase expression in third instar larval discs (Figure IV-6). The β-galactosidase expression was seen in the anterior compartment of all observable imaginal discs in both the heterozygous control group and in larvae homozygous for the vg^{21r7} chromosome (Figure IV-6A and B). Eye color is consistently red in the vg^{21r7} heterozygous flies (Figure IV-6C). The presence of two copies of the vg^{21r7} chromosome reduces expression of the w^{+} gene as seen by a decrease or absence of eye color (Figure IV-6D). To ensure that the flies which appeared to have no expression of *white*$^{+}$ in the eye still retained the P[*lacZ w*$^{+}$]*ci* insert, adults were stained for β-galactosidase expression which could be observed in the adult antennae (Figure IV-6D). It is clear that

while the expression of *white*$^{+}$ in the presence of the four P[21-3] elements on chromosome II is reduced or lost entirely, β-galactosidase expression remains consistent. I conclude that the additional *P* elements cause a strong repression effect at the additional *white* locus. The results from the analysis of the response of the wing phenotype to repressor suggested that each additional element in vg^{21r7} contributes to the repression of vg^{21-3} as there is a higher degree of wing malformation with the loss of each element.

Next, I tested if the P[*lacZ w*$^{+}$] insert also showed altered repression in the presence of the recombinant stocks. As Figure IV-7 shows, we tested the P[*lacZ w*$^{+}$]*ci* insert with $vg^{21r7\text{-}FR}$, $vg^{21r7\text{-}L\text{-}FR}$, vg^{21r36} and vg^{21-3}. In contrast to the wild type colored eye shown in Figure IV-7A, weak repression is evident in the $vg^{21r7\text{-}FR}$ with significant lightening at the eye periphery (Figure IV-7B). The $vg^{21r7\text{-}L\text{-}FR}$ stock produces mild repression resulting in an orange eye color (Figure IV-7C). Somewhat surprisingly, repression of the *w*$^{+}$ insert is seen in stocks with only a single *P* element as in the vg^{21r36} that produces a light orange eye shown in Figure IV-7D and vg^{21-3} that produces a dark orange eye in Figure IV-7E. This demonstrates that the P[21r36] and P[21-3] elements can produce *P* repressor; the next question is what proteins are produced by these elements and how do they compare to Type I and Type II repressors.

Coding analysis

It is evident from the wing repression tests in Chapter III and presented in this chapter, and the *white*$^{+}$ repression tests described above that P[21r36] and P[21-3] elements exhibit some properties of somatic *P* repressors. By looking at the sequences present in these elements, we sought to determine if either correspond to a canonical

Type I or Type II element. Simulated translations of these *P* elements based on sequence obtained previously (Chapter III) are used to compare the encoded proteins to known repressors in Figure IV-8. In Figure IV-8A, P[21r36] and P[21] proteins are aligned with that of a KP (Type II) *P* element. Protein domain searching reveals that P[21r36] contains a DNA binding domain, highlighted in blue in Figure IV-8A, called the THAP domain. P[21] encodes an incomplete THAP domain (ROUSSIGNE *et al.* 2003) that diverges from the consensus in the last 7 amino acids and would be predicted not to exhibit any repressive qualities. The replacement of the hydrophobic valine at position 77 with a hydrophilic lysine also suggests the remaining THAP sequence would not assemble properly as it is a highly conserved residue present in the fourth helical region of the protein (BESSIERE *et al.* 2007). The P[21r36] protein is different from the KP element as it is missing the leucine zipper from amino acids 101-122, which functions in the dimerization of repressor (Figure IV-8A, orange highlighted sequence) (ANDREWS and GLOOR 1995; LEE *et al.* 1996). This would suggest that DNA binding, but not dimerization, is sufficient for some repressor activity. In a comparison of the P[21-3] encoded repressor to the minimal Type I repressor (Figure IV-8B), we see that they are identical for 532 amino acids at which point the P[21-3] repressor continues on for 10 novel amino acids, and the minimal Type I for 30 amino acids. Both proteins contain full THAP and leucine zipper domains which are utilized in repression.

Insulated P[21-3] repression

The apparent expression of sufficient repressor protein from a P[21-3] element at 3 different genomic locations to alter the vg^{21-3} wing phenotype leads us to question what

allows for repressor production in the wing disc where it would be required. Either we have been fortunate to find 3 locations which all allow for repressor accumulation in the wing to alter the wing phenotype, or the *P* element can express itself in the wing without any external enhancement. To test which of these alternatives is more likely we constructed a nested *P* element where the P[21-3] element was inserted in an inverted orientation within a pH-Stinger (BAROLO *et al.* 2000) element. The pH-Stinger element contains gypsy insulators surrounding the multiple cloning site to minimize the effect of external enhancers (Figure IV-9A, B). Without inverting the internal element, a <5'-<5'-3'>-3'> orientation could give not only the desired <5'-<5'-3'>-3'> insertion product (Figure IV-9A, line 1), but could also produce a <5'-<5'-3'> insertion, a <5'-3'>-3'> insertion or a <5'-3'> insertion (Figure IV-9A, lines 2-4). The disadvantage is the insertion of a product shown in Figure IV-9A, line 4, which would retain the w^+ marker to allow for selection, but would be uninsulated at the 3' end. With the internal element inverted with respect to the external element in a <5'-<3'-5'>-3'> fashion, the double element could only insert in one of two ways: first, to give the desired <5'-<3'-5'>-3'> element or second, just the <3'-5'> internal element (Figure IV-9B). Of those, only the desired product with both insulators would also contain the *white*$^+$ marker (Figure IV-9B, line 1). Out of approximately 200 G0 flies, only a single w^+ transformant was produced, illustrating the rarity with which transposase cuts the outer, but not inner, element. The pH-Stinger-P[21-3] insert was found to be homozygous lethal on the third chromosome and was used to produce $vg^{21\text{-}3}$; pH-Stinger-P[21-3]/+ flies. The wings from these flies were indistinguishable from $vg^{21\text{-}3}$ flies as can be seen by comparing panels C and D of

Figure IV-9. This leads us to conclude that a *P* element which is insulated from genomic enhancers does not produce sufficient repressor to alter the wing phenotype of $vg^{21\text{-}3}$.

Discussion

By using two somatic tests for repression: the effect of repressor on *vg* insert wing phenotype and on P[lacZ w^+] expression, it has been demonstrated that the products which are encoded by P[21-3] and P[21r36] elements can act as *P* element repressors. While these two *P* elements are able to repress the eye color rescue from P[*lacZ* w^+]*ci* they do not exhibit variegation of the eye color which indicates that the mechanism of repression may be different from the repression seen by KP elements (Scott Hanna, personal communication). The variegation from a KP repressor can also be modified by certain chromatin modifiers so it is a possibility that the KP protein is able to interact with additional proteins which are not recruited to the repression site by the *P* elements used in this study.

It is interesting to note the repression seen by P[21-3] elements acting on $vg^{21\text{-}3}$ is both additive in nature and can be highly variable. If we take the average wing phenotype of a homozygous vg^{21r7} population as a wild type score of 5, when the same stock is placed over a null allele it exhibits some wing malformation with 6.8 as the average score (Figure IV-5, blue bars). The removal of the rightmost element in $vg^{21r7\text{-}FR}$ further increases the heterozygous phenotype average to 13 (Figure IV-5, green bars) and removing both outer elements in $vg^{21r7\text{-}L\text{-}FR}$ reduces the heterozygous phenotype average

to 5.4 (Figure IV-5, red bars). The nearly wild type population from $vg^{21r7\text{-}L\text{-}FR}$ suggests that the single additional repressor-making element at 2R50B1 can restore the wing phenotype better when the elements at 2L33E7 and 2R60B4 are not present. This may be indicative that not only are the additional elements interacting to repress at the *vg* locus, but they also would be competing to repress one another. The true contribution of each element to repression of $vg^{21\text{-}3}$ can only be accurately measured by observing the phenotypes produced when only the *P* at *vg* and a single repressor element are present. The isolation of each element to observe with $vg^{21\text{-}3}$ is ongoing.

Due to the weak nature of the *P* element promoter, as evidenced by its use in enhancer traps (O'KANE and GEHRING 1987), it is highly prone to position-effect differences in expression levels. Evidence pointing to the insertion-site specific variations in repressor production was obtained when we utilized the insulated pH-Stinger-P[21-3] insert. The insulated P[21-3] element is less prone to external position effects and would allow us to test repressor production driven by *P* element sequences in isolation. This experiment gave wings which were comparable in size to $vg^{21\text{-}3}$ alone (Figure IV-9), which supports the proposal that expression driven by the *P* element is not sufficient to produce repressor in the developing wing. As each of the new P[21-3] insertion sites has been found to contribute to repression of the $vg^{21\text{-}3}$ wing phenotype, each must be expressed under the influence of an enhancer which allows for repressor accumulation in the wing disc cell lineage. As P[21-3]60B4 has now been isolated from the other *P* elements in the initial vg^{21r7} stock and appears to contribute to repression it is a good candidate for further studies into how this element's expression profile is modified by external genomic factors.

This chapter documents two *P* elements that act as *P* repressors, but do not fall into the Type I or Type II categories. Clearly, the categories were somewhat arbitrary since they do not fully encompass all repressors made by natural *P* elements. The difference between a P[21] element which likely does not make repressor and a P[21r36] element at the same genomic location which does produce repressor is that P[21r36] contains 16 more amino acids in common with a KP repressor, and of those 7 are at the C-terminal end of the THAP domain. The THAP domain is a DNA binding domain present in both vertebrate and invertebrate proteins (CLOUAIRE *et al.* 2005). The THAP domain also corresponds to the area of the KP element encoded repressor which is responsible for DNA binding (LEE *et al.* 1996). The minimum protein required for *P* element binding and repression is likely confined to the THAP domain, but the ability of this protein to repress is certainly enhanced by inclusion of the leucine zipper dimerization domain (LEE *et al.* 1996), and possibly by further unstudied sequences present in the full length *P* element repressor which may aid in cellular localization, folding or stability. Another factor which contributes strongly to the level of repression of an individual element, but which is not encoded by the element itself, is the genomic location and expression of that *P* element. This is important to consider as a weak repressor protein which is highly expressed may function better than a strong repressor protein which does not get produced as the *P* element is inserted into a silenced region of chromatin.

A more useful categorization system for *P* element repressors might be to designate which domains are present. The three initial regions would be the THAP domain (THAP), the leucine zipper (LZIP), and the *P* element somatic inhibitor (PSI)

binding site located near the end of the third exon (ADAMS *et al.* 1997). The P[21r36] element would then be designated a THAP *P* element, KP and P[21-3] elements would be THAP-LZIP *P* elements and any former Type I repressor, which by the previous size definition does retain the PSI site, would be categorized as a THAP-LZIP-PSI *P* element. Should additional functional domains be discovered in the future, they could be easily incorporated into this scheme. *P* elements can be given these designations based on sequence alone without any need for repressor testing as no claim is made to the strength of repressor or amount of repressor made, although experimental evidence does clearly suggest that the more domains present will produce a stronger repressor. As the ability to make repressor is closely tied to the insertion site of the element and this repression may work for some tests or tissues but not others, it is much more difficult to come up with a standard designation for repressor ability than repressor potential. A workable solution could utilize a system where each test of repression is listed as a single letter superscript to the type of repressor. If the standard tests are done in bristles (sn^{weak} mutability), eyes (w^{+} repression), germline (gonadal dysgenesis), and wings (vg^{21-3} suppression) then the letters b, e, g, and w could be used. An example would be the vg^{21-3} element, designated P[21-3]*vg* THAP-LZIP $^{e+w-}$. This indicates the name, location, repressor domains and the results of repressor tests in a single and clearly defined notation. A functionally defined notation for *P* element repressors is needed to keep track of the increasing number of repressors which have been studied to better enable researchers to predict the properties of *P* elements and stocks which carry them.

It is necessary to note that the results from P[21r36] indicate moderate repressor activity from a very small repressor element containing the complete THAP domain. As

many common *P* element vectors and enhancer traps contain more *P* element sequence than P[21r36], they can encode a protein with a full THAP domain. The elements could then give unexpected results if researchers were not aware that self-repression by THAP-containing elements has now been demonstrated. The majority of elements which are based on pCaSpeR vectors could be prone to self-repression as they contain the first 586 bases of a normal P element (THUMMEL *et al.* 1988). The opinion that commonly used and largely internally deleted *P* elements such as pCaSpeR "almost certainly does not encode a repressor polypeptide" (RONSSERAY *et al.* 2001) must be questioned in order to better explain the repressive qualities exhibited by these and other small *P* elements. The retained *P* element sequences in these derivatives include the protein coding region for the first 125 amino acids of *P* repressor which makes them not only 32 amino acids longer than the product of P[21r36], but these pCaSpeR elements are THAP-LZIP elements and have the potential to bind *P* element DNA as dimers. The P[lacZ w^+] elements which are based on pCaSpeR have β-galactosidase fused to this THAP-LZIP protein which may greatly impair the ability to form proper repressor. Still, P[lacZ w^+] elements have been shown multiple times to cause odd *P* repression effects which could be at least partially explained by the inclusion of a THAP domain in the *P* encoded transcript (LEMAITRE and COEN 1991; ROCHE and RIO 1998; RONSSERAY *et al.* 2001).

Figure IV-1: Inverse PCR screening of *vg*^{21r7} *P* elements. A. The P[21-3] element is shown with the positions of Sau3A I restriction sites. Digestion with Sau3A I yields a 5' end fragment of 1397 bp plus the length of genomic DNA to the nearest Sau3A I site. The fragment is depicted above the P[21-3] sequence with primer sites used marked in red and green. **B.** Digested genomic DNA is ligated in conditions which promote circularization. **C.** PCR using the KAN 11 (red) and KAN 12 (green) primers amplifies the inverted region between the two arrows shown in B. The nested PELEM 2 primer (magenta) is then used for sequencing of products obtained by gel extraction of the inverse PCR bands shown with arrows in **D**. Bands extracted are the result of KAN11/12 amplification since single primer products were undetectable.

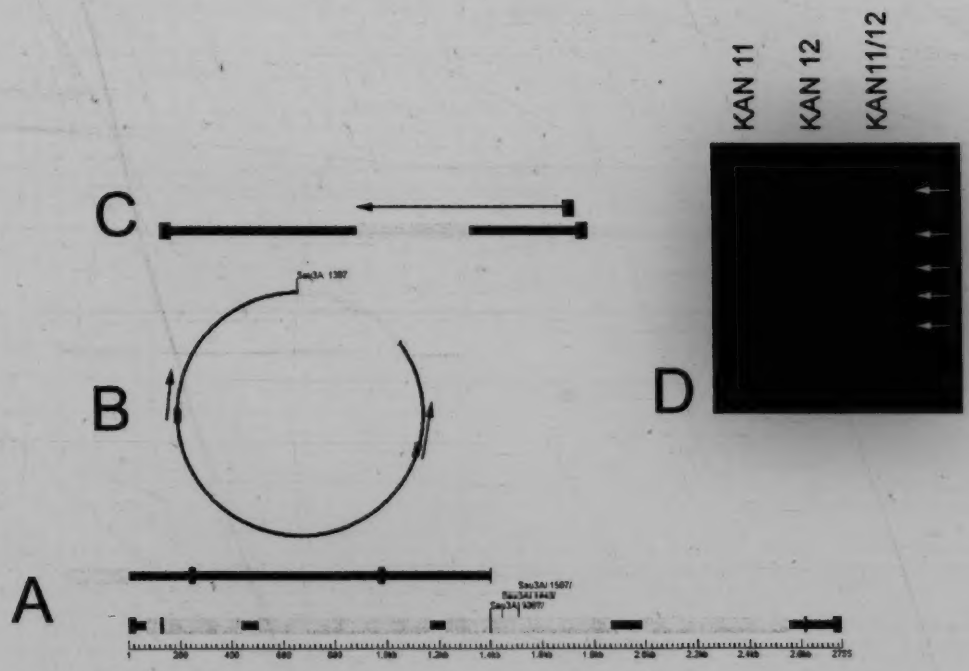

KAN 11
KAN 12
KAN11/12
C
B
A
D

Figure IV-2: Alignment of recombinant markers and vg^{21r7} *P* elements on chromosome 2. The second chromosome of *Drosophila melanogaster* is depicted with the left arm shown in blue and the right arm shown in green with the scale beneath showing distance in megabase pairs (M). **A.** The position of the mutations b^1, pr^1, c^1, px^1, sp^1 used to select for recombinants is shown above the location of the 4 *P* elements in the vg^{21r7} stock in **B.** The elements are as follows: green = P[21-3]33E7, red = P[21-3] at vestigial, blue = P[21-3]50B1 and orange = P[21-3]60B4. The recombinant stocks obtained from selection are designated: **C.** $vg^{21r7\text{-}FR}$ **D.** $vg^{21r7\text{-}L\text{-}FR}$ and **E.** P[21-3]60B4.

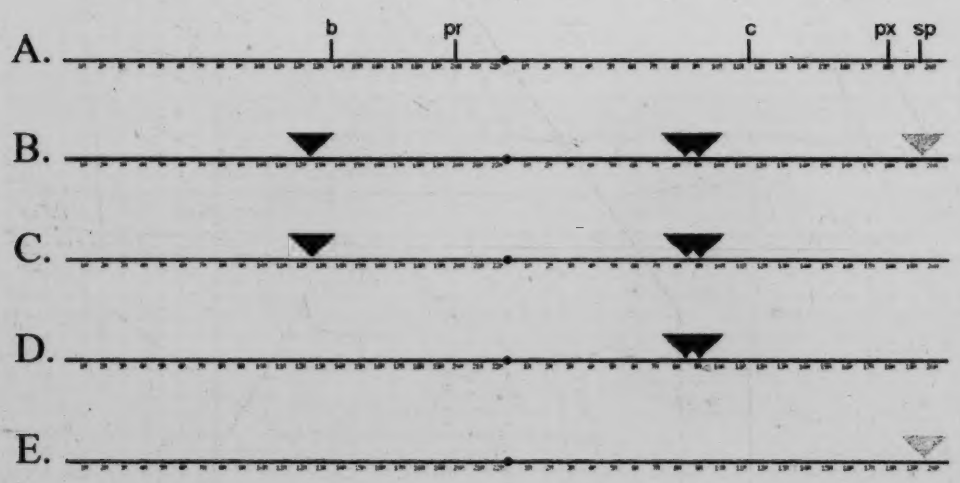
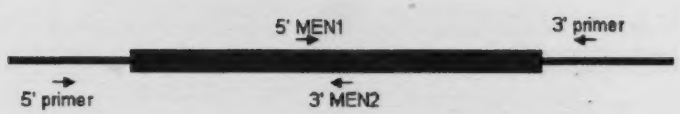

A.
b
pr
c
px
sp
B.
C.
D.
E.

Figure IV-3: PCR primers used to study new insertions. The standard PCR used to amplify each new insertion utilized a 5' genomic primer in combination with MEN 2 and a 3' genomic primer in combination with MEN 1. **A.** The general location of the primers used is shown. **B.** The sequence of each genomic primer used for the 3 new insertion sites and the MEN 1 and MEN 2 are shown.

A



B

Insertion Site	Primer	Sequence
P[21-3]33E7	5' – KAN13	cgacgcgcaggaaaaccaaataga
	3' – KAN14	gccagcggatatgttgcaatttgtg
P[21-3]50B1	5' – KAN 15	attccgatgcgagcgcgactg
	3' – KAN 16B	cccgttagcagcacccccacac
P[21-3]60B4	5' – KAN 17	aagcgagcacggcgtaaaaatcaatg
	3' – KAN 18	cggtcgcagggctgggctattac
P element internal	5' – MEN 1	caccgctagctcgatcagacgc
	3' – MEN 2	cgcagtttcggtggcattttcaa

Figure IV-4: Southern analysis of vg^{21r7} recombinant stocks. Lane designations are as follows: **A** – vg^{21r7} **B** – $vg^{21r7\text{-}FR}$ **C** – $vg^{21r7\text{-}L\text{-}FR}$ and **D** – P[21-3]2R60. The numbers shown are the sizes in base pairs predicted from P[21-3] inserts at the known genomic locations. The high molecular weight band seen in all lanes is present in every *P* element *vg* stock studied.

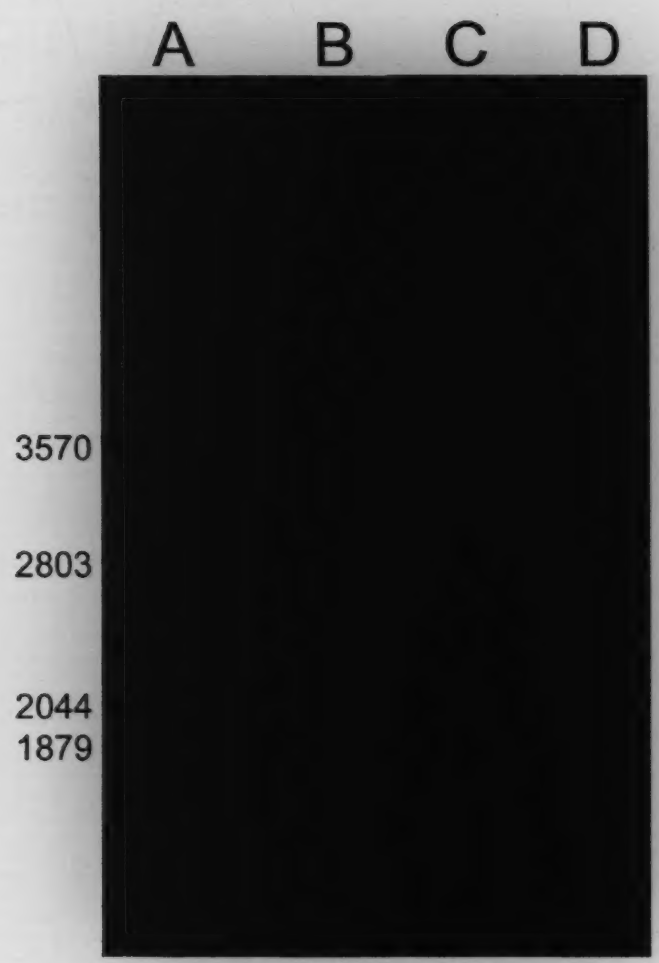

A
B
C
D
3570
2803
2044
1879

Figure IV-5: Phenotype of vg^{21r7} recombinants over vg^{B} at 25°C. Females of the vg^{21r7} stock and 2 recombinant stocks which retain the P[21-3] insertion at *vg* were crossed to male vg^{B}/*CyO* flies and the progeny were scored in heterozygous combination with vg^{B} for wing phenotype. The height of each bar corresponds to the number of wings scored while the phenotype observed is shown along the x-axis. A numerical scale of wing size (Figure I-3) is shown below the illustration of wing shapes. The vg^{21r7}/vg^{B} combination is shown in blue and ranges from wild type to 10. The $vg^{21r7\text{-}FR}/vg^{B}$ combination is shown in green and ranges from wild type to 21. The $vg^{21r7\text{-}L\text{-}FR}/vg^{B}$ combination is shown in red and ranges from wild type to 10.

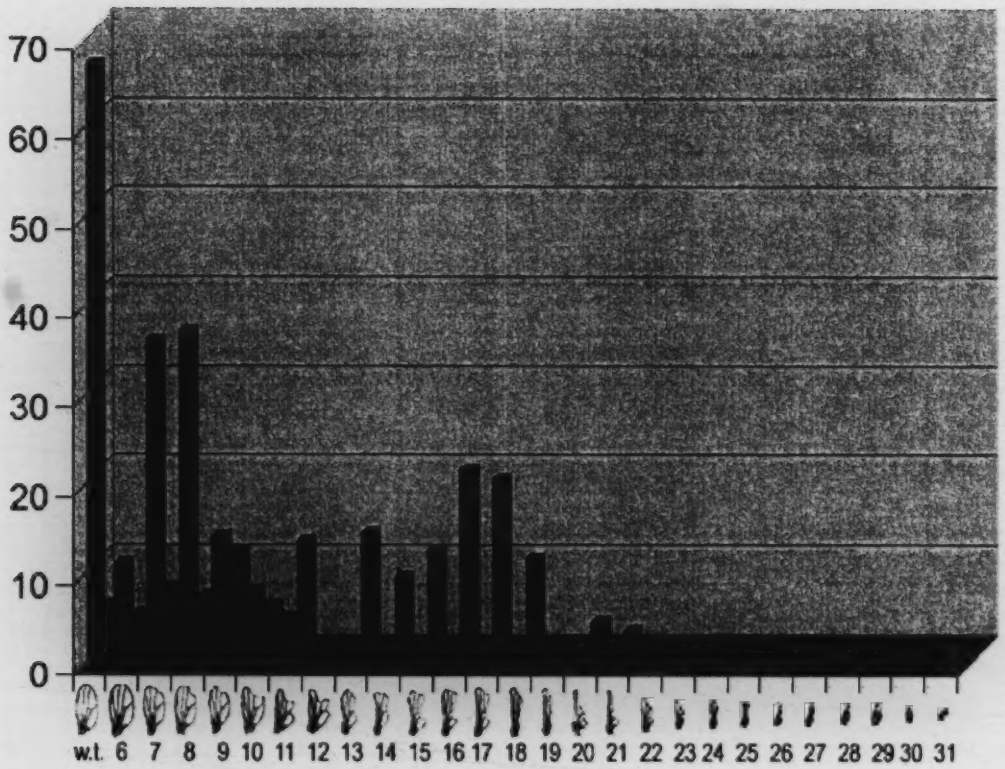

70
60
50
40
30
20
10
0
w.t. 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

Figure IV-6: Expression of P[*lacZ* w^+] markers in combination with vg^{21r7}. Wing and leg imaginal disc staining for β-galactosidase in **A.** heterozygous vg^{21r7}/+; P[*lacZ* w^+] *ci*/+ and **B.** homozygous vg^{21r7}; P[*lacZ* w^+] *ci*/+. The range of eye color phenotypes in adults also stained for β-galactosidase is shown for the heterozygotes in **C.** and homozygotes in **D.**

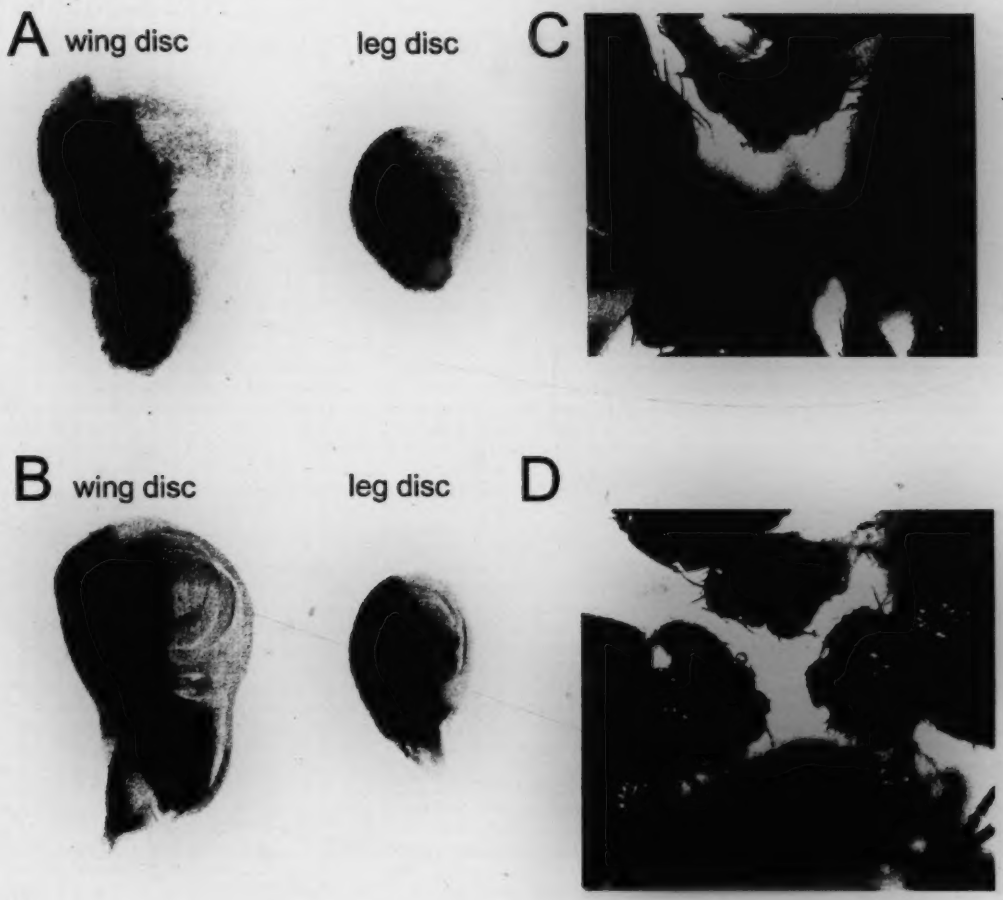

A wing disc
leg disc
C
B wing disc
leg disc
D

Figure IV-7: P[*lacZ* w^+]*ci* expression in stocks carrying *P* repressor. Eye color was consistent among individuals of each genotype with an indicative eye shown. The phenotype observed when a single copy of P[*lacZ* w^+]*ci* is placed with a homozygous: **A.** wild type second chromosome **B.** $vg^{21r7\text{-}FR}$ **C.** $vg^{21r7\text{-}L\text{-}FR}$ **D.** vg^{21r36} and **E.** $vg^{21\text{-}3}$.

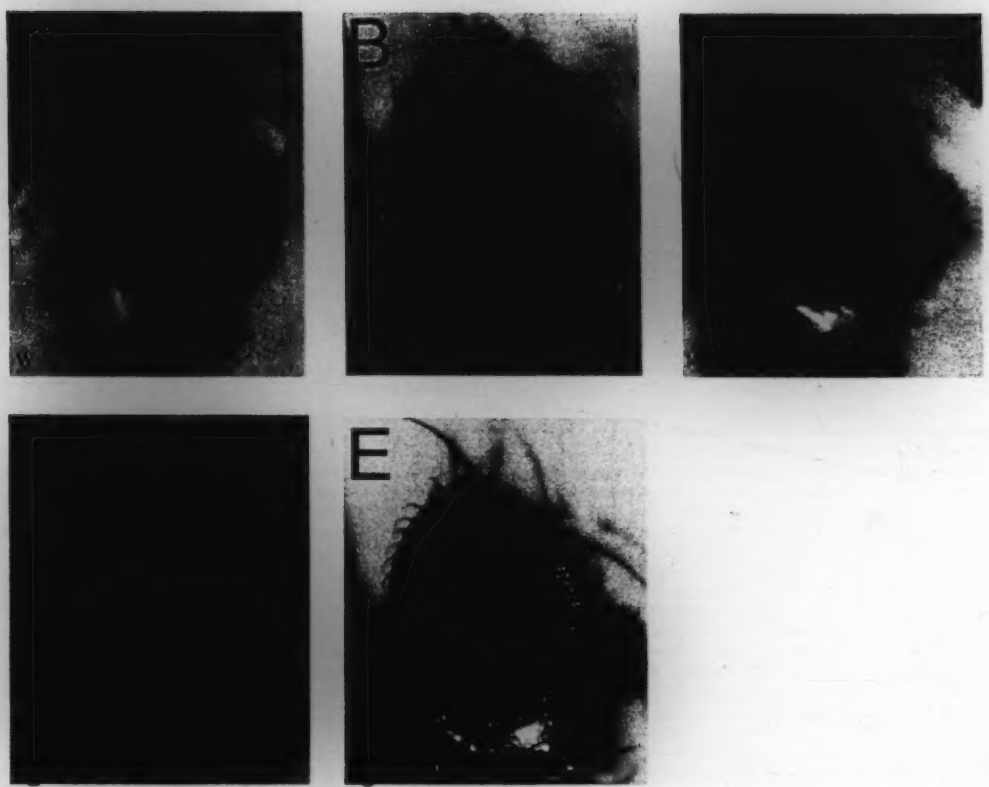

B
E

Figure IV-8: Coding capacity of new repressors. A. The predicted proteins encoded by a KP element Type II repressor, P[21r36] and P[21] are aligned. The location where the P[21] protein diverges from the KP is marked as a purple arrow and the location where the P[21r36] protein diverges is marked by a red arrow. The THAP domain is marked in blue and the leucine zipper marked in orange. **B.** The proteins encoded by the minimal Type I repressor (*P* element truncated at base 1950) and a P[21-3] element are aligned. The site of sequence divergence is marked with a blue arrow.

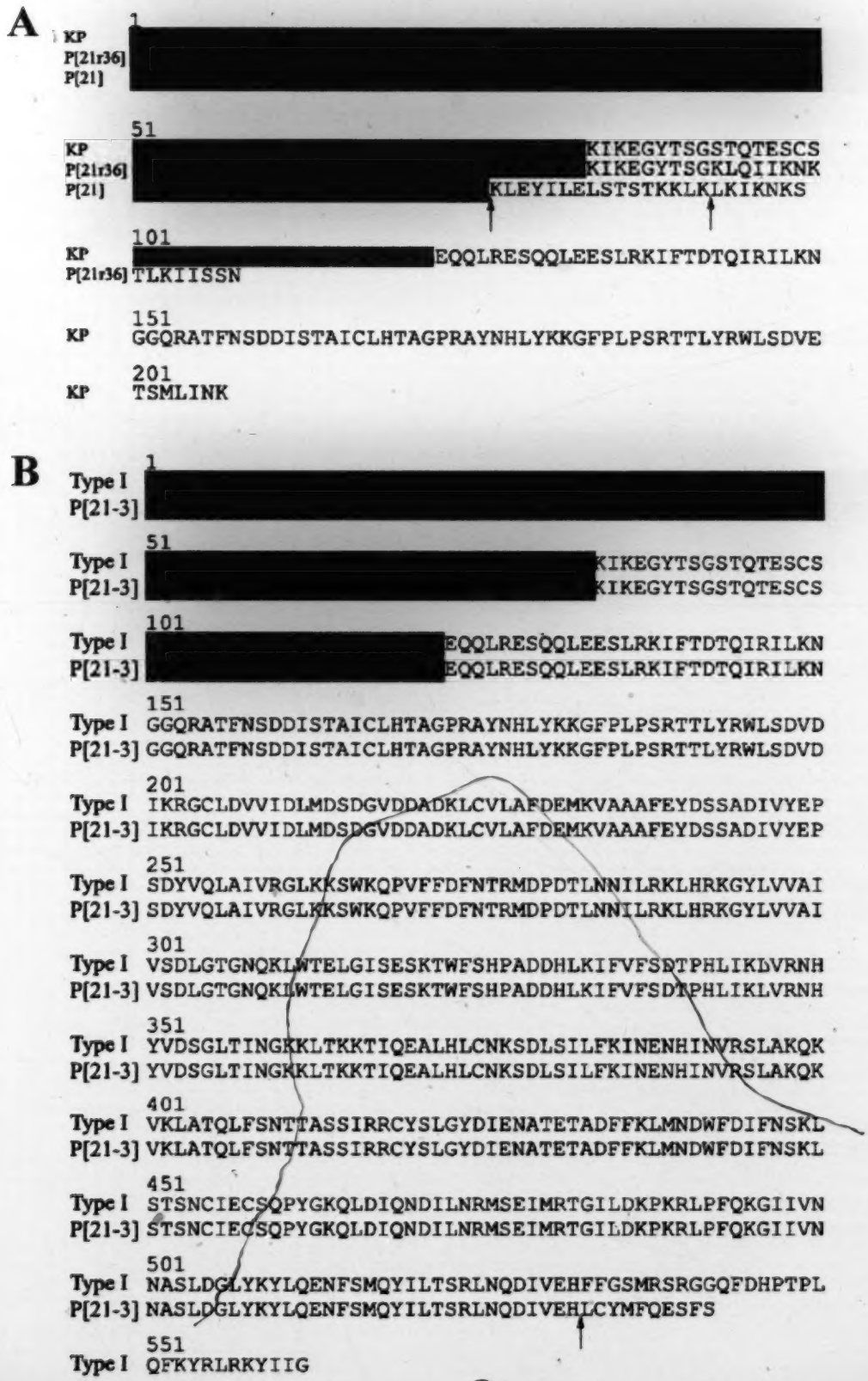
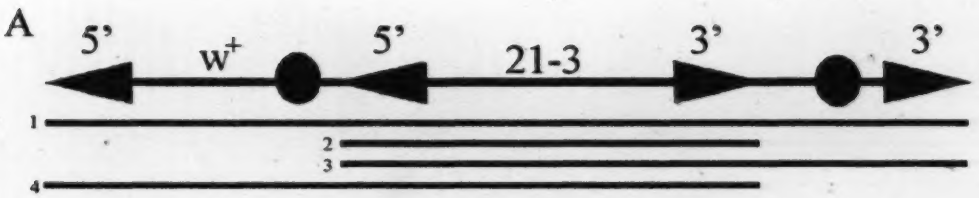
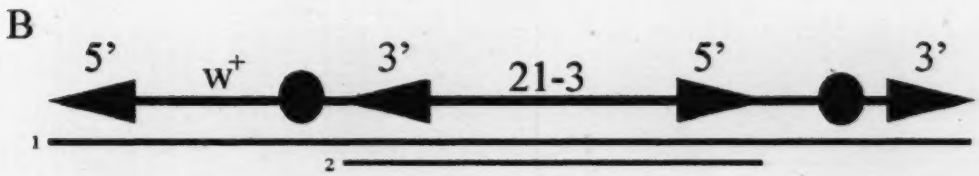
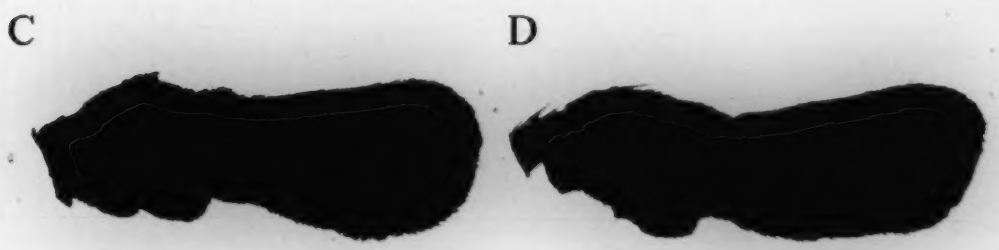

A
KP
P[21r36]
P[21]
1
51
KIKEGYTSGSTQTESCS
KIKEGYTSGKLQIIKNK
KLEYILELSTSTKKLKLKIKNKS
101
EQQLRESQQLEESLRKIFTDTQIRILKN
TLKIISSN
151
GGQRATFNSDDISTAICLHTAGPRAYNHLYKKGFPLPSRTTLYRWLSDVE
201
TSMLINK
B
Type I
P[21-3]
1
51
KIKEGYTSGSTQTESCS
KIKEGYTSGSTQTESCS
101
EQQLRESQQLEESLRKIFTDTQIRILKN
EQQLRESQQLEESLRKIFTDTQIRILKN
151
GGQRATFNSDDISTAICLHTAGPRAYNHLYKKGFPLPSRTTLYRWLSDVD
GGQRATFNSDDISTAICLHTAGPRAYNHLYKKGFPLPSRTTLYRWLSDVD
201
IKRGCLDVVIDLMDSDGVDDADKLCVLAFDEMKVAAAFEYDSSADIVYEP
IKRGCLDVVIDLMDSDGVDDADKLCVLAFDEMKVAAAFEYDSSADIVYEP
251
SDYVQLAIVRGLKKSWKQPVFFDFNTRMDPDTLNNILRKLHRKGYLVVAI
SDYVQLAIVRGLKKSWKQPVFFDFNTRMDPDTLNNILRKLHRKGYLVVAI
301
VSDLGTGNQKLWTELGISESKTWFSHPADDHLKIFVFSDTPHLIKLVRNH
VSDLGTGNQKLWTELGISESKTWFSHPADDHLKIFVFSDTPHLIKLVRNH
351
YVDSGLTINGKKLTKKTIQEALHLCNKSDLSILFKINENHINVRSLAKQK
YVDSGLTINGKKLTKKTIQEALHLCNKSDLSILFKINENHINVRSLAKQK
401
VKLATQLFSNTTASSIRRCYSLGYDIENATETADFFKLMNDWFDIFNSKL
VKLATQLFSNTTASSIRRCYSLGYDIENATETADFFKLMNDWFDIFNSKL
451
STSNCIECSQPYGKQLDIQNDILNRMSEIMRTGILDKPKRLPFQKGIIVN
STSNCIECSQPYGKQLDIQNDILNRMSEIMRTGILDKPKRLPFQKGIIVN
501
NASLDGLYKYLQENFSMQYILTSRLNQDIVEHFFGSMRSRGGQFDHPTPL
NASLDGLYKYLQENFSMQYILTSRLNQDIVEHLCYMFQESFS
551
QFKYRLRKYIIG

Figure IV-9: Repression test of an insulated P[21-3]. The 5' ends of each *P* element are marked in green while the 3' ends are marked in blue. The gypsy insulators are the circles surrounding the P[21-3] element. **A.** The four lines under the construct represent the possible mobilization products in a non-inverted arrangement. **B.** The two lines under the construct represent the possible mobilization products using an inverted internal element. **C.** A representative wing of a *w*; $vg^{21\text{-}3}$ fly. **D.** A representative wing of a *w*; $vg^{21\text{-}3}$; +/ pH-Stinger-P[21-3] fly.







w; *vg*$^{21\text{-}3}$; +

w; *vg*$^{21\text{-}3}$; + / pH-Stinger-P[21-3]

Literature Cited

ADAMS, M. D., R. S. TARNG and D. C. RIO, 1997 The alternative splicing factor PSI regulates P-element third intron splicing in vivo. Genes Dev **11:** 129-138.

ANDERSON, K. J., M. M. DAVIS and R. B. HODGETTS, 2006 Novel events associated with phenotypic reversion of a P element mutant in Drosophila melanogaster. Genome **49:** 1184-1192.

ANDREWS, J. D., and G. B. GLOOR, 1995 A role for the KP leucine zipper in regulating P element transposition in Drosophila melanogaster. Genetics **141:** 587-594.

BAROLO, S., L. A. CARVER and J. W. POSAKONY, 2000 GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in Drosophila. Biotechniques **29:** 726, 728, 730, 732.

BESSIERE, D., C. LACROIX, S. CAMPAGNE, V. ECOCHARD, V. GUILLET *et al.*, 2007 Structure-function analysis of the thap-zinc finger of thap1, a large C2CH DNA-binding module linked to RB/E2F pathways. J Biol Chem.

BLACK, D. M., M. S. JACKSON, M. G. KIDWELL and G. A. DOVER, 1987 KP elements repress P-induced hybrid dysgenesis in Drosophila melanogaster. Embo J **6:** 4125-4135.

BUSHEY, D., and J. LOCKE, 2004 Mutations in Su(var)205 and Su(var)3-7 suppress P-element-dependent silencing in Drosophila melanogaster. Genetics **168:** 1395-1411.

CLOUAIRE, T., M. ROUSSIGNE, V. ECOCHARD, C. MATHE, F. AMALRIC *et al.*, 2005 The THAP domain of THAP1 is a large C2CH module with zinc-dependent sequence-specific DNA-binding activity. Proc Natl Acad Sci U S A **102:** 6907-6912.

ENGELS, W. R., 1996 P Elements in Drosophila, pp.

GLOOR, G. B., C. R. PRESTON, D. M. JOHNSON-SCHLITZ, N. A. NASSIF, R. W. PHILLIS *et al.*, 1993 Type I repressors of P element mobility. Genetics **135:** 81-95.

KIM, J., A. SEBRING, J. J. ESCH, M. E. KRAUS, K. VORWERK *et al.*, 1996 Integration of positional signals and regulation of wing formation and identity by Drosophila vestigial gene. Nature **382:** 133-138.

LEE, C. C., Y. M. MUL and D. C. RIO, 1996 The Drosophila P-element KP repressor protein dimerizes and interacts with multiple sites on P-element DNA. Mol Cell Biol **16:** 5616-5622.

LEMAITRE, B., and D. COEN, 1991 P regulatory products repress in vivo the P promoter activity in P-lacZ fusion genes. Proc Natl Acad Sci U S A **88:** 4419-4423.

NITASAKA, E., and T. YAMAZAKI, 1988 Excision of one of two defective P elements as the cause of alternate mutational events (sn+ and sne) of the singed bristle allele snw in Drosophila melanogaster. Jpn J Genet **63:** 303-312.

O'KANE, C. J., and W. J. GEHRING, 1987 Detection in situ of genomic regulatory elements in Drosophila. Proc Natl Acad Sci U S A **84:** 9123-9127.

RASMUSSON, K. E., J. D. RAYMOND and M. J. SIMMONS, 1993 Repression of hybrid dysgenesis in Drosophila melanogaster by individual naturally occurring P elements. Genetics **133:** 605-622.

REHM, E. J., 2003 Inverse PCR & Cycle Sequencing of P Element Insertions for STS Generation, pp. Berkeley Drosophila Genome Project.

ROBERTSON, H. M., and W. R. ENGELS, 1989 Modified P elements that mimic the P cytotype in Drosophila melanogaster. Genetics **123:** 815-824.

ROCHE, S. E., and D. C. RIO, 1998 Trans-silencing by P elements inserted in subtelomeric heterochromatin involves the Drosophila Polycomb group gene, Enhancer of zeste. Genetics **149:** 1839-1855.

RONSSERAY, S., A. BOIVIN and D. ANXOLABEHERE, 2001 P-Element repression in Drosophila melanogaster by variegating clusters of P-lacZ-white transgenes. Genetics **159:** 1631-1642.

ROUSSIGNE, M., S. KOSSIDA, A. C. LAVIGNE, T. CLOUAIRE, V. ECOCHARD *et al.*, 2003 The THAP domain: a novel protein motif with similarity to the DNA-binding domain of P element transposase. Trends Biochem Sci **28:** 66-69.

THUMMEL, C. S., A. M. BOULET and H. D. LIPSHITZ, 1988 Vectors for Drosophila P-element-mediated transformation and tissue culture transfection. Gene **74:** 445-456.

WILLIAMS, J. A., S. S. PAPPU and J. B. BELL, 1988 Molecular analysis of hybrid dysgenesis-induced derivatives of a P-element allele at the vg locus. Mol Cell Biol **8:** 1489-1497.

Chapter V: The suppressible vg^{21-3} phenotype is caused by a microRNA produced by *P* element transcription.

Introduction

It is rare to have the opportunity to observe rapid genomic change in an animal species while it is occurring, rather than just observing the differences between closely related species. The horizontal transmission of *P* elements to *Drosophila melanogaster*, likely from *Drosophila willistoni*, has given us just that opportunity (DANIELS *et al.* 1990). In the last 50 years, *P* elements have gone from undetectable in *D. melanogaster* to become ubiquitous in all genomes worldwide except for populations maintained in laboratories (KIDWELL 1983). The rapid invasion of these elements caused a phenomenon termed hybrid dysgenesis in which male strains carrying *P* elements (P strains) mated to female strains without *P* elements (M strains) created progeny with high rates of sterility, pupal lethality and male recombination (KIDWELL *et al.* 1977). These phenotypes result from the insertion of *P* elements into new genomic locations causing insertional mutations and genomic rearrangements created by the double strand breaks that arise during mobilization (GEYER *et al.* 1988; KAUFMAN and RIO 1992; NITASAKA and YAMAZAKI 1988; STAVELEY *et al.* 1995). Reciprocal crosses are spared the adverse effects of *P* element transposition since elements that are passed on through a female produce a cytoplasmic repressor (termed P cytotype), which prevents the expression of the enzyme transposase that is required for mobilization (KIDWELL 1981; LASKI *et al.* 1986). The creation of new alleles by *P* element insertion mutagenesis using both natural and modified elements has been an invaluable tool for *Drosophila* geneticists, but in some instances the mechanism of how *P* element insertion impairs the gene is not adequately understood. A number of *P* element insertions of *vestigial* (*vg*) fall into this

class as they cause a mutant wing phenotype that is reverted in the presence of P repressor protein (ANDERSON *et al.* 2006; WILLIAMS *et al.* 1988b).

P elements inserted at *vg* are a useful tool because wing area yields a quantifiable measure of *P* repressor activity. The mutant $vg^{21\text{-}3}$ produces an extreme wing phenotype that can be suppressed to give wild type wings in the presence of a repressor-making *P* element (ANDERSON *et al.* 2006; HODGETTS and O'KEEFE 2001; WILLIAMS *et al.* 1988a). Despite its usefulness as a repressor test (GLOOR *et al.* 1993), the exact nature of why insertion of the P[21-3] element into the *vg* 5' UTR produces a strong phenotype that is suppressed by P repressor is unclear. Experiments to address this have been carried out (HODGETTS and O'KEEFE 2001) and the authors proposed a model based on *P* element transcriptional interference. There are several other insertions of *P* elements at the same nucleotide position as P[21-3] that act very differently in terms of wing phenotype severity or response to *P* repressor. These may be useful in determining what is responsible for allele severity and repressor suppression. The vg^{21} allele was the first *P* element *vg* mutant produced and it was subsequently utilized in cloning the gene (WILLIAMS and BELL 1988). The allele was a cryptic mutation that produced nicked wings in combination with a null allele, but when the hemizygote was combined with a *P* repressor, normal wings were observed. From this original mutant, several secondary *P* element replacements have been created that lie at the same genomic location and in the same orientation as the original P[21] element. These include the severe mutant $vg^{21\text{-}3}$, which is also responsive to *P* element repressor like vg^{21}, and the mild mutant vg^{2a33}, which does not respond to P repressor (HODGETTS and O'KEEFE 2001; WILLIAMS *et al.* 1988a). By comparing the phenotypes of these different alleles under different situations

in the context of the underlying *P* element sequence, we aim to provide a mechanistic explanation of the mutant effects of the *vestigial P* element alleles.

At the molecular level, all the *P* element *vg* alleles have an insertion 423 bases downstream from the *vestigial* transcriptional start site (HODGETTS and O'KEEFE 2001). All the insertions are transcribed collinearly with the *vg* transcript such that both the *vg* and *P* element promoter are transcribed in the same direction. The *P* element terminator has been demonstrated to allow transcription to pass through and into the remainder of the *vg* gene, which allows for the production of any functional VG product (HODGETTS and O'KEEFE 2001). As the phenotypic difference between alleles cannot be explained by insertion site or orientation differences in the *P* elements, it must be due to the sequences contained within the *P* element. By looking at numerous alleles it has been observed that deletion of the *P* element promoter in some alleles produces reversions (STAVELEY *et al.* 1994), which supports the theory that *P* element transcription can be deleterious to *vg* function.

P element repressor is able to downregulate the expression of *lacZ* fusion genes driven by the *P* promoter (LEMAITRE and COEN 1991), which suggests that the repression of wing phenotype in alleles such as $vg^{21\text{-}3}$ may be due to the prevention of transcription from the *P* promoter. The non-repressible allele, vg^{2a33}, contains a 46 bp deletion that removes one of the *P* repressor binding sites located within the 3' 11 bp inverted repeat. This adds support to the theory that suppression of *vestigial P* element alleles is due to transcriptional repression (HODGETTS and O'KEEFE 2001). These authors further postulated that the relative level of *vg*-initiated to *P*-initiated transcripts was responsible for the severity of the wing phenotype observed. By using some modified nuclease

protection assays, I have attempted to measure the amount of transcription occurring at *vg* and *P* promoters and to determine whether the transcription at the *P* promoter is reduced in the presence of a single strong *P* element repressor, P[Sal]89D.

If *P* transcription is deleterious to *vg* transcript function, it may be acting through the RNA silencing pathway. RNA silencing can be endogenously encoded as microRNA (miRNA) products or by transgenic short interfering RNAs (siRNA) (BARTEL 2004; HAMILTON and BAULCOMBE 1999). While they retain some of the same machinery to silence genes post-transcriptionally, miRNA leads to the inhibition of translation while siRNA leads to mRNA degradation (for review, see TANG 2005). In this chapter, I address the possibility that a *P* element at *vg* can create a mutant wing phenotype through RNA silencing by determining the effects of mutations in components of the RNA induced silencing complex (RISC) on the wing phenotype.

Results

Dual transcript nuclease protection assay

To test the theory (HODGETTS and O'KEEFE 2001) that the allele vg^{21-3} is more severe than vg^{21} due to an increase in the ratio of *P* element to *vg* initiated transcription, we needed to be able to compare the amount of transcript produced from each promoter in a way which was internally standardized. To accomplish this, an oligonucleotide-based nuclease protection assay was developed. This utilized a single 48 base DNA fragment designed to permit the visualization of protected fragments initiated at either promoter while still retaining a non-complementary 3' end to differentiate between the undigested oligonucleotide from the products protected by the upstream *vg* initiated

transcript (Figure V-1A). The 48 nucleotide fragment spanned the *P* transcriptional start site with 30 bases downstream, and 11 bases upstream. An additional 7 non-complementary bases were added on to the upstream end to distinguish the low level of *vg* transcript from the undigested radiolabeled oligonucleotide, present in excess. Experimental conditions that allowed for the detection of *vg* and *P* initiated transcripts in the same reaction were optimized, but the production of band staggering due to slight instability in the free ends of the otherwise properly base paired products could not be eliminated. The fragment intensities were quantified, including the nearby staggered products, and the relative amounts of *P* to *vg* transcription were measured (Figure V-1B). In vg^{21} pupae the amount of RNA coming from the *P* element promoter was found to be 28 times more abundant than transcripts coming from the *vg* promoter by quantifying signal intensity of the main band recorded by the phosphorimager. Similarly, in vg^{21-3} it was found that the *P* element initiated transcripts were 30 times more abundant than the *vg* promoter initiated transcript. The less than 10% difference in relative transcript levels is unlikely to be responsible for the large difference in wing phenotype observed between the wild type winged vg^{21} individuals and the absent winged vg^{21-3} individuals.

_P_ transcription at _vg_ with repressor present

While the difference in phenotype between vg^{21} and vg^{21-3} could not be readily explained by differences in levels of *P* element transcription, the nuclease protection assay was used to determine the amount of *P* element transcription in stocks with *P* element repressor present versus those without repressor in an attempt to explain the suppressive effects of *P* repressor on the vg^{21-3} phenotype. The mutant allele was

combined with a single *P* element situated on chromosome III, which has previously been shown to modify the phenotype of some *vg* alleles (ROBERTSON and ENGELS 1989). Designated P[Sal]89D, the element is a modified Pc[ry^{+}] element that contains a premature stop codon in the fourth exon at a Sal I restriction site. This premature stop codon prevents the element from making transposase but does not alter the production of full length 66 kD repressor protein (KARESS and RUBIN 1984). The element is inserted on chromosome III at polytene band position 89D and is a homozygous lethal insertion. The stocks *w*; $vg^{21\text{-}3}$; P[Sal]89D *Sb* / TM2 *Ubx* and vg^{2a33}; P[Sal]89D *Sb* / TM2 *Ubx* were made and *P* initiated transcription was compared to that of the parental stocks not carrying P[Sal]89D. As $vg^{21\text{-}3}$ is known to be a *P* repressible stock, the phenotype, as expected, was restored to normal. On the other hand, vg^{2a33}, which does not respond to *P* repressor, exhibited a moderate mutant wing phenotype. Figure V-2A shows the results of the nuclease protection assay and the corresponding RNA loading control slot blot in Figure V-2B. The same 48 oligonucleotide used to compare *P* to *vg* transcription was used in this assay, but the focus is solely on the 30 nt fragment protected by *P* initiated transcription. Obviously, the 30 nt protected fragment can arise from both the *P* elements at *vg* and the single element at 89D, when present. After correcting for variations in RNA loaded from each sample, the intensity of *P* transcription was compared between the samples. Several experiments were averaged to analyze the amount of *P* element transcription in our stocks. It was found that the addition of P[Sal]89D to $vg^{21\text{-}3}$ reduced the transcripts to 62%, while it reduced *P* transcripts in vg^{2a33} to 52% of the amount seen without repressor. As the P[Sal]89D element cannot be creating a negative amount of transcription, it must be concluded that each of the *P* elements at *vg* is downregulated by

the *P* element repressor produced by P[Sal]89D as the total transcription from all three elements is less than that arising from just the two elements at *vg*. While these results were reproducible, we were unable to discriminate the *P* element at *vestigial* initiated transcripts from those initiating at P[Sal]89D to get a precise measure of the decrease in transcription at *vg* alone.

RNA interference

There is ample evidence from our repressor results and previous studies (HODGETTS and O'KEEFE 2001) that transcription from *P* elements within the *vg* 5' UTR is somehow deleterious to VG function. A reduction in the amount of *P* element transcript through deletion of the *P* element promoter or addition of *P* element repressor is restorative to wing development (HODGETTS and O'KEEFE 2001; STAVELEY *et al.* 1994; WILLIAMS *et al.* 1988b). However, results from the nuclease protection assays comparing *P* element initiated transcription in vg^{21} and $vg^{21\text{-}3}$ suggest that the difference in phenotype among these two alleles with wild type *P* element promoters is not caused by the small transcriptional differences, but may be due the the different sequences contained within the elements. An alternative hypothesis is that the transcripts themselves somehow interact to result in lowered VG production. There is extensive predicted secondary structure within the 5' UTR of the *vg* mRNA which could allow for an interaction between the *vg* and *P* initiated transcripts (HODGETTS and O'KEEFE 2001). If the two RNA products do interact they may trigger the silencing of functional *vg* mRNA through the RNA-induced silencing complex (RISC). As P[Sal]89D was shown to lower the amount of *P* element transcription for both $vg^{21\text{-}3}$ and vg^{2a33} but only restored

the phenotype of $vg^{21\text{-}3}$, it can be argued that the *P* element transcript is deleterious in $vg^{21\text{-}3}$ mutants while the *vg* transcript is fully functional. On the other hand, for vg^{2a33}, the modified *vg* transcript may produce the wing phenotype. A prediction of the mRNA secondary structure may yield clues as to why the *P* element mRNA might contribute to the mutant wing phenotype in $vg^{21\text{-}3}$ but not vg^{2a33}. The sequences of the two *P* element transcripts were folded using the online RNA secondary structure prediction program MFOLD (ZUKER *et al.* 1999) which determines dozens of thermodynamically favorable secondary structures allowing for common folding regions to be identified. The 2377 base mRNA predicted for P[21-3] spans the sequence between the *P* element promoter (HODGETTS and O'KEEFE 2001) to *P* element terminator, with the introns removed. The same rationale is used for the 418 base P[2a33] mRNA, except there are no introns to remove in the sequence as the deletion extends from the first to last exon. The P[21-3] mRNA has extensive stem loop secondary structure with one particular region selected for further analysis (Figure V-3A and C). This region was isolated by taking bases 301 to 350 and bases 1751 to 1800 to make a single continuous 100 bp sequence (Figure V-3C). This fragment was used to query a microRNA prediction program called MiRScan (http://genes.mit.edu/mirscan/) which compares the predicted RNA secondary structure to that of known pre-miRNA structures and ranks the query based on likelihood of producing a miRNA. The MiRScan program found a hypothetical microRNA from bases 1757 to 1777 of the P[21-3] mRNA (Figure V-3C). This putative microRNA scored 7.624 which is in the top 2% of scores for stem loops in the large samples used to predict microRNAs in published human and worm screens (LIM *et al.* 2003a; LIM *et al.* 2003b). Although this score is promising, it does rank in the bottom 5% of known microRNAs.

The P[2a33] mRNA sequence also had many stem loops present, but none with even the 21 bp minimum length needed to provide a microRNA (Figure V-3B).

If the sequences in the *P* element initiated transcript are involved in mediating the severity of the wing phenotype seen in vg^{21-3} through the production of a microRNA, then mutants in components of the RNA-induced silencing complex (RISC) should alleviate the wing phenotype. To minimize confounding factors, RISC mutants had to be selected which were not carrying *P* elements. A thorough search yielded *spindle* E^1 and $armitage^{72.1}$ as candidate mutants that were RISC components and located on chromosome III, which would facilitate study in combination with vg^{21-3}. Mutant stocks were obtained from the Bloomington *Drosophila* Stock Center and used to create vg^{21-3}; $spnE^1$/TM2 *Ubx e* and vg^{21-3}; $armi^{72.1}$/TM2 *Ubx e* stocks. As the phenotype of *P* element *vestigial* mutants is generally less severe at 18°C than 25°C or 29°C multiple temperatures were used to test the new stocks. The control stock of vg^1; $spnE^1$/TM2 *Ubx e* was tested only at 29°C as this is the temperature for maximal wing growth and lower temperatures resulted in absent wings. Figure V-4 shows the wing sizes measured while Figure V-5 depicts representative wings from each cross. While vg^1 wings were identical in the presence or absence of $spnE^1$, the vg^{21-3} flies demonstrated between 1.4 and 1.7 times larger wing area with $spnE^1$ than with the wild type control. The $armi^{72.1}$ mutant produced 1.5 times larger wings in combination with vg^{21-3} at 25°C, but wings of comparable size when the temperature was reduced to 18°C.

These RISC results prompted a search for additional mutants to test, and since few third chromosome genes in the RNA silencing pathway have been identified, our search expanded to include second chromosome genes. One candidate, and a key

component of RISC, is *Argonaute 1* (*Ago1*), which is less then 1 Mbp away from *vestigial*. This initiated a search for deletions that may span both genes and the stock Df(2R)CX1, wg^{12} b^{1} pr^{1}/ SM1 was found to be the only one available. Comparisons between $vg^{21\text{-}3}$ and vg^{2a33} as homozygotes or in combination with the deletions Df(2R)vg^{B}, which does not span *Ago1*, and Df(2R)CX1, which does span *Ago1*, are shown in Figure V-6. The wing phenotype in $vg^{21\text{-}3}$/ Df(2R)CX1 was restored much more than with any other RISC mutant tested (Figure V-6 A, B, C) while vg^{2a33}/Df(2R)CX1 was a severe phenotype similar to when Df(2R)vg^{B} was used (Figure V-6 D,E,F).

Discussion

The phenotype of $vg^{21\text{-}3}$ is unique in that it can produce wings which are merely stumps, and yet in the presence of repressor protein encoded elsewhere in the genome, the wings may be restored to wild type (ROBERTSON and ENGELS 1989; WILLIAMS *et al.* 1988b). By comparing the P[21-3] element to other elements which are located at the same genomic location, such as P[21] and P[2a33], which have very different phenotypic qualities, we can better understand the basis of both wing impairment and restoration from *P* elements. While different alleles have been used in this study, the focus is on understanding $vg^{21\text{-}3}$.

The few *P* element alleles of *vestigial* to be studied that have a deletion for the *P* element promoter have a wild type or cryptic phenotype, suggesting that transcription from the *P* element may contribute to the severity of the wing phenotype (ANDERSON *et al.* 2006; WILLIAMS *et al.* 1988a). The two theories proposed to explain this observation

(HODGETTS and O'KEEFE 2001), involve transcriptional blockage as the *vg* initiated RNA polymerase complex may not pass through the *P* promoter properly, or that levels of *P* transcription interfere with VG production at a post-transcriptional level. The finding that *vg* transcripts read through the *P* promoter without a decrease in transcript levels makes the hypothesis of post-transcriptional interference more likely than a transcriptional block (HODGETTS and O'KEEFE 2001). Our nuclease protection assay was designed to detect the relative levels of *P* promoted and *vg* promoted transcription and we used it to compare the ratios of a the weak allele vg^{21} to the strong mutant $vg^{21\text{-}3}$. The ratios detected were 28:1 and 30:1 for vg^{21} and $vg^{21\text{-}3}$ respectively. While there is a slight increase in the amount of relative *P* element transcription in the more severe allele, it is unlikely that a 7% variation in ratios is the sole factor that determines the wide range of wing phenotypes produced from these mutants.

If it is not the amount of transcription from the *P* element promoter which is the main cause of phenotypic differences between alleles, then it may be the sequence within the transcript that affects the phenotype. In order to add support to the theory that the transcripts themselves are deleterious to *vg* function, we sought to compare the amount of *P* element transcription in the presence of *P* element repressor protein to the *P* element transcription without repressor present. When we compare two alleles with very similar *P* element sequence, vg^{21} and its derivative vg^{2a33}, one main difference is the inability of *P* repressor to suppress the vg^{2a33} wing phenotype. The only molecular difference between the P[2a33] and P[21] elements is that P[2a33] contains a 46 bp deletion that removes the 3' 11 bp repeat, one of the 6 repressor binding sites (HODGETTS and O'KEEFE 2001; LEE *et al.* 1996). From this information, our initial hypothesis was that

while *P* repressor may lower *P* element transcription in some alleles to suppress them, loss of repressor binding sites will prevent this transcriptional downregulation and the phenotype will remain. There have been studies done on fusion *P* elements that contain LacZ driven by the *P* element promoter, which have shown substantial reduction in *β-galatosidase* staining and loss of LacZ mRNA in all cell types studied when mated to P cytotype flies (LEMAITRE and COEN 1991; LEMAITRE *et al.* 1993) These studies, however, did not assess transcription from endogenous or unmodified *P* elements. This prompted our assay of *P* element transcription levels in two *vg* alleles. To measure the amount of *P* element initiated transcription, we used the nuclease protection assay on the two mutants $vg^{21\text{-}3}$ and vg^{2a33} expecting that while we would see a reduction in P[21-3] transcription with P[Sal]89D present, the level of P[2a33] transcription would not decrease. To our surprise, the amount of *P* element initiated transcription decreased substantially in both stocks, indicating that the repressor produced by P[Sal]89D is able to bind both P[21-3] and P[2a33] to lower *P* transcription.

The finding that the amount of *P* element transcript correlated with the $vg^{21\text{-}3}$ suppression, but not vg^{2a33}, led to a reassessment of what may be causing the wing phenotype in each case. Since vg^{2a33} has a moderate wing phenotype regardless of *P* element transcription, then it may be simply a case of the *P* element sequence in the 5' untranslated region of the *vg* mRNA impairing the stability or translatability of the message. Because $vg^{21\text{-}3}$ flies can go from a strong mutant phenotype to wild type wings when the transcription of the *P* element is repressed, then the *vg* initiated transcript must be functional and it is the *P* transcript which is deleterious. One possible explanation is that the P[21-3] transcript worked through the *Drosophila* RNA silencing pathway to

accomplish the reduction in VG expression. Post-transcriptional gene silencing relies on extended double-stranded regions of RNA as the catalyst to promote mRNA destruction or silencing, so the predicted transcripts from P[21-3] and P[2a33] were used to model the secondary structure each would assume at physiological conditions. In comparing the 2377 nucleotide P[21-3] mRNA to the 418 nucleotide P[2a33] mRNA structures it was observed that the P[21-3] transcript contained a much longer double stranded region with no side branches than P[2a33], which contained no unbranched regions of more than 20 base pairs. This extended stem-loop in P[21-3] could potentially be used to promote RNA silencing and the sequence of this region was scored for similarity to known microRNA structures using the MiRScan program. This sequence of bases 301 to 350 and 1751 to 1800 yielded a hypothetical 21 nucleotide microRNA from 1757 to 1777 which scored above the 98th percentile for microRNA structure homology when compared to stem-loop structures from other animal genomes. Although the theoretical finding that vg^{21-3} flies may have a severe wing phenotype due to mRNA silencing from microRNAs was promising, we still needed experimental evidence to substantiate the hypothesis. By hindering the ability of the vg^{21-3} flies to produce microRNAs and assaying the wing phenotype, we could firmly prove or disprove the RNA silencing hypothesis.

Because *P* elements that produce repressor can modify the phenotype of vg^{21-3} (GLOOR *et al.* 1993; WILLIAMS *et al.* 1988b), the search for available RNA silencing mutants was limited to genes not on chromosome II with *vg*, and not created by *P* element mutagenesis themselves. These constraints restricted our analysis to mutants in two genes involved in RNAi. The first mutant tested, *spindle* E^1 is defective in the

production of a DEAD-box RNA helicase and has been shown to cause a decrease in gene silencing when heterozygous with a wild type allele (KENNERDELL *et al.* 2002; PAL-BHADRA *et al.* 2004). The second mutant tested, *armitage*$^{72.1}$, is a non-DEAD-box RNA helicase created by excision of an enhancer trap element which produced a deletion in the 5' UTR of *armitage*, but does not disrupt the coding region of the gene (TOMARI *et al.* 2004). This mutant does not prevent the production of ~22nt silencing dsRNAs, but it blocks the step of unwinding of the dsRNA to allow for RISC assembly (TOMARI *et al.* 2004; ZAMBON *et al.* 2006). Careful observation of wings at various temperatures from *vg*$^{21-3}$ in combination with the two RNAi mutants showed that in each case, wing size was increased by approximately 50% when the production of RISC is impaired. The conflicting result that *armitage*$^{72.1}$ was effective at restoring the wing phenotype at 25°C, but not at 18°C, might be due to a temperature-sensitive effect on the folding of the partially deleted 5' UTR of *armitage* which prevents translation only at higher temperatures.

Members of the Argonaute family are central proteins of the RNA-induced silencing complex. They will produce mRNA silencing when an imperfectly paired small RNA is present in RISC, or mRNA cleavage when a fully complementary small RNA is present (SCHWARZ *et al.* 2002). In *Drosophila*, it has been found that there are multiple Argonaute proteins with distinct functions in the silencing pathway. *Argonaute 1* (*Ago1*) is required for miRNA maturation and is dispensable for siRNA mediated mRNA cleavage, while *Argonaute 2* (*Ago2*) is required for siRNA assembly into RISC (OKAMURA *et al.* 2004). We tested *vg*$^{21-3}$ in combination with a deletion which spanned both *vg* and *Ago1* loci and found almost complete wing restoration. The *vg*^{2a33} allele was

also tested with the deletion and the resulting wings were unaltered in phenotype. It is an enticing possibility that the *P* element transcript arising in $vg^{21\text{-}3}$ inhibits VG function via an RNA silencing mechanism that specifically involves the microRNA pathway and will be the focus of future experiments. A model of the likely pathway of $vg^{21\text{-}3}$ transcriptional silencing is shown in Figure V-7.

While certainly not common, the discovery of *P* element insertions which cause mutant phenotypes which can be alleviated by the presence of P repressor or a P cytotype is not confined to *vestigial* alleles (ROBERTSON and ENGELS 1989). There are likely more, but the effort of testing the phenotype in the presence of properly expressed P repressor is of little interest to non-P element researchers. Regardless of the number or frequency of suppressible *P* element mutations, it is important to note that transposon insertion can cause a non-permanent modification of the surrounding gene's function through the RNA silencing pathway. If this mode of gene expression modification is present in wild populations, the fitness advantage to retaining repressor producing elements extends past the prevention of further transposition to silencing existing mutations.

Figure V-1: Nuclease protection assay to detect *vg* and *P* initiated transcripts. A. The sequence of the *P* element is shown with the *vg* promoted transcripts initiating 488 bp upstream of the nucleotide labeled Pp, which is the +1 site for *P* element transcription. The 48 nucleotide oligonucleotide used is shown directly below the genomic sequence with the non-complementary bases shown in bold and tilted away from the genomic sequence. The predicted fragments produced in the nuclease protection assay would be the undigested 48 nucleotide fragment, a 41 nucleotide fragment produced from protection by binding the *vg* initiated transcript and a 30 nucleotide fragment which would be produced by binding to the *P* element initiated transcript. **B.** Results from the nuclease protection assay show the predicted 48, 41 and 30 nucleotide fragments derived from RNA extracted from vg^{21} and $vg^{21\text{-}3}$ pupae. The staggered products around the 41 and 30 nucleotide fragments are due to base-pair instability at fragment ends as evidenced by conducting the experiment at different digestion temperatures.

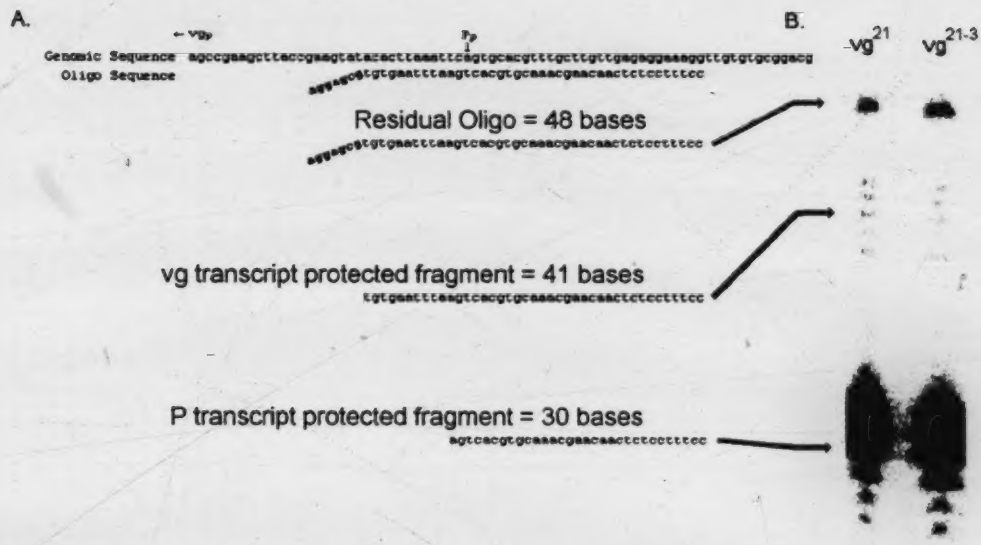

A.
B.
Vg21
Vg21-3
Genomic Sequence
Oligo Sequence
Residual Oligo = 48 bases
vg transcript protected fragment = 41 bases
P transcript protected fragment = 30 bases

Figure V-2: Nuclease protection assay to compare *P* initiated transcription with and without repressor. **A.** The 48 base oligonucleotide (Figure V-1A) was used to compare RNA preparations from vg^{21-3} and vg^{2a33} pupae to vg^{21-3} and vg^{2a33} pupae which also carried P[Sal]89D. **B.** An amount of RNA equal to that used for the nuclease protection was used for a slot blot. The blot was probed with a DNA fragment from the ribosomal protein gene, RpL32, to allow cross sample comparisons.

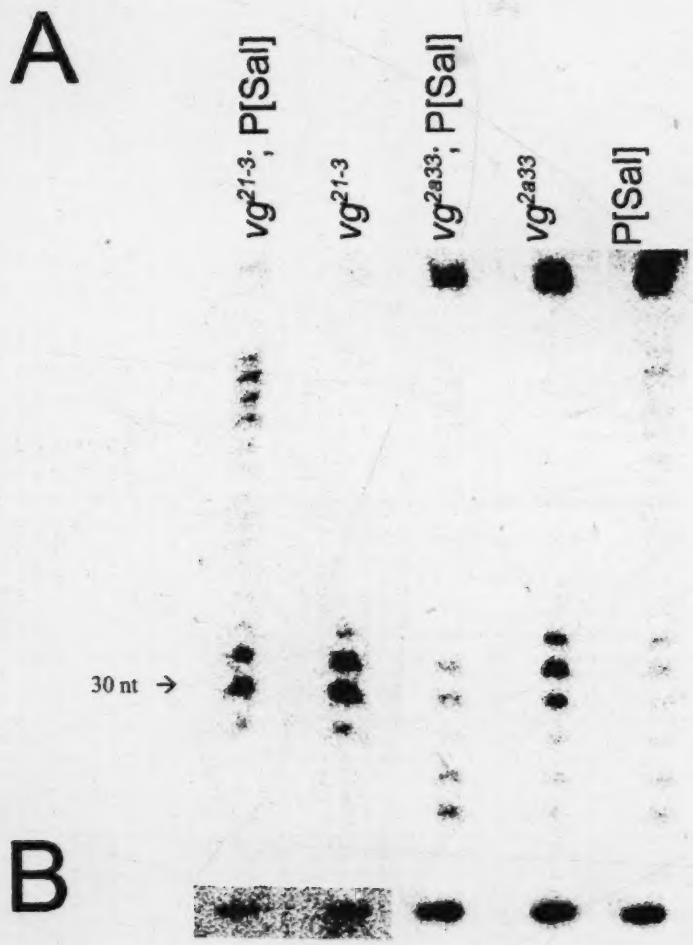

A
vg21-3; P[Sal]
vg21-3
vg2a33; P[Sal]
vg2a33
P[Sal]
30 nt →
B

Figure V-3: *P* element transcript predicted secondary structure. A. The structure of the predicted 2377 base mRNA transcript of P[21-3]. The area marked in red is represented at a larger scale in part **C**. **B.** The structure of the predicted 418 base mRNA transcript of P[2a33]. **C.** The long dsRNA region from the P[21-3] transcript with a predicted microRNA sequence marked in green.

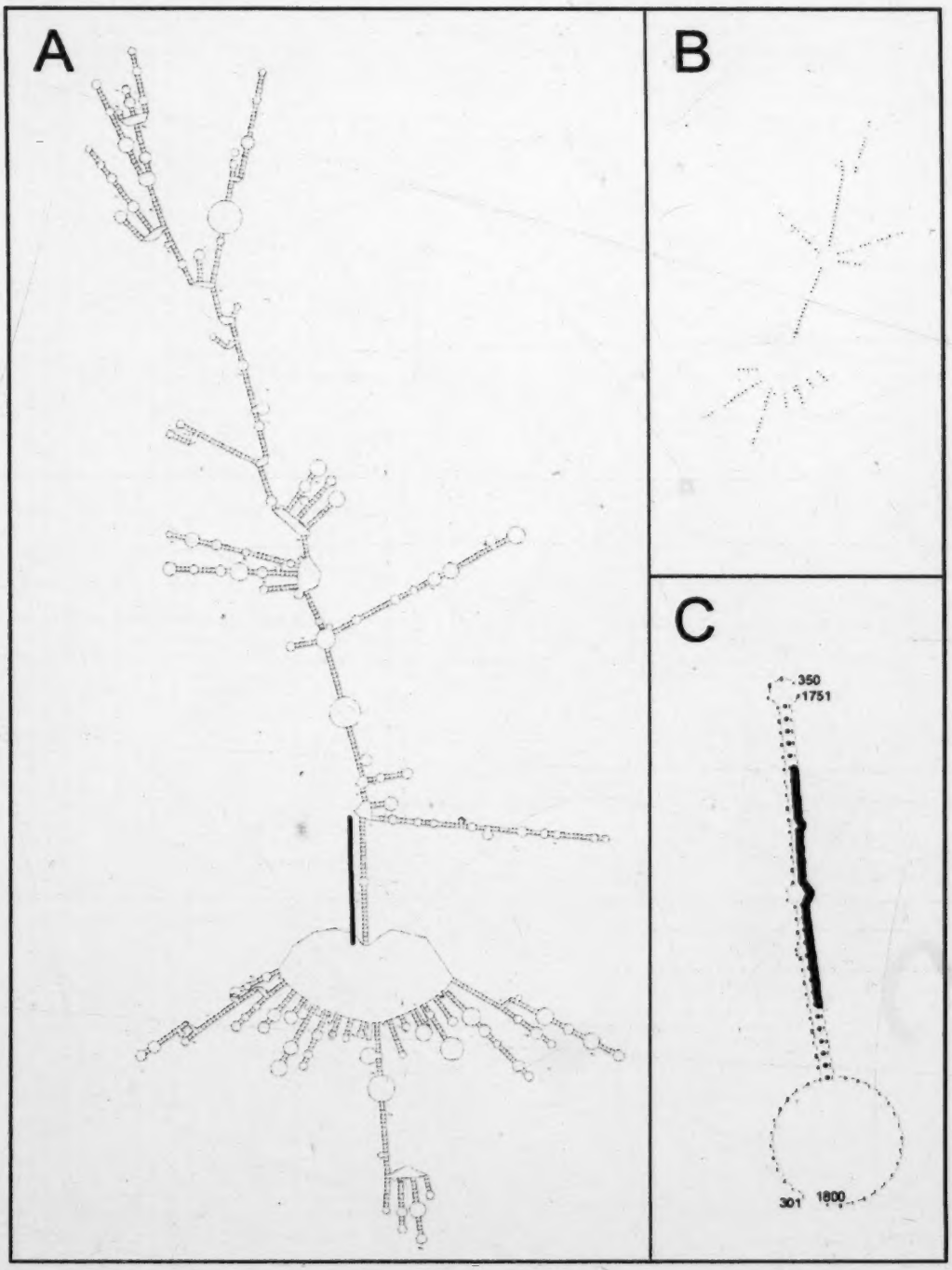

A
B
C
350
1751
301
1800

Figure V-4: Wing area for *vg* mutants in combination with RISC mutants. Flies tested were homozygous for the *vg* allele and heterozygous for the RISC allele shown (or a wild type chromosome) and the balancer TM2 *Ubx e*. Wings were dissected, fixed, and then photographed which allowed total wing area to be measured digitally. The average area of 20 dissected wings for every sample is represented with 95% confidence interval error bars shown.

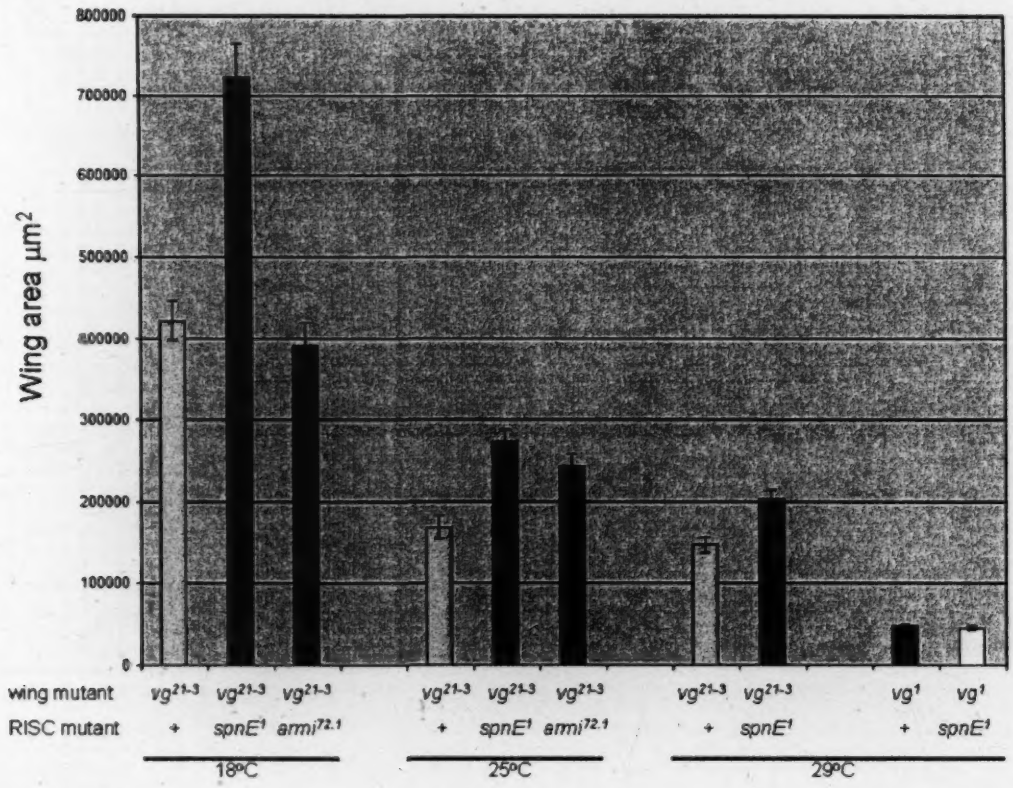

800000
700000
600000
500000
400000
300000
200000
100000
0
Wing area µm²
wing mutant
RISC mutant
vg21-3 vg21-3 vg21-3 vg21-3 vg21-3 vg21-3 vg21-3 vg21-3 vg1 vg1
+ spnE1 armi72.1 + spnE1 armi72.1 + spnE1 + spnE1
18°C
25°C
29°C

Figure V-5: Representative wings from *vg* mutants in combination with RISC mutants. A wing from each stock matching the mean wing size is shown for comparison. **(A, B, C)** are from flies grown at 18°C **(D, E, F)** are from flies grown at 25°C and **(G, H, I, J)** are from flies grown at 29°C. **(A, D, G)** vg^{21-3} ; +/TM2 *Ubx e* **(B, E, H)** vg^{21-3}; $spnE^{1}$/TM2 *Ubx e* **(C, F)** vg^{21-3}; $armi^{72.1}$/TM2 *Ubx e* **(I)** vg^{1}; +/TM2 *Ubx e* **(J)** vg^{1}; $spnE^{1}$/TM2 *Ubx e*. Statistically significant variations in wing size are seen in **B, E, F** and **H**.

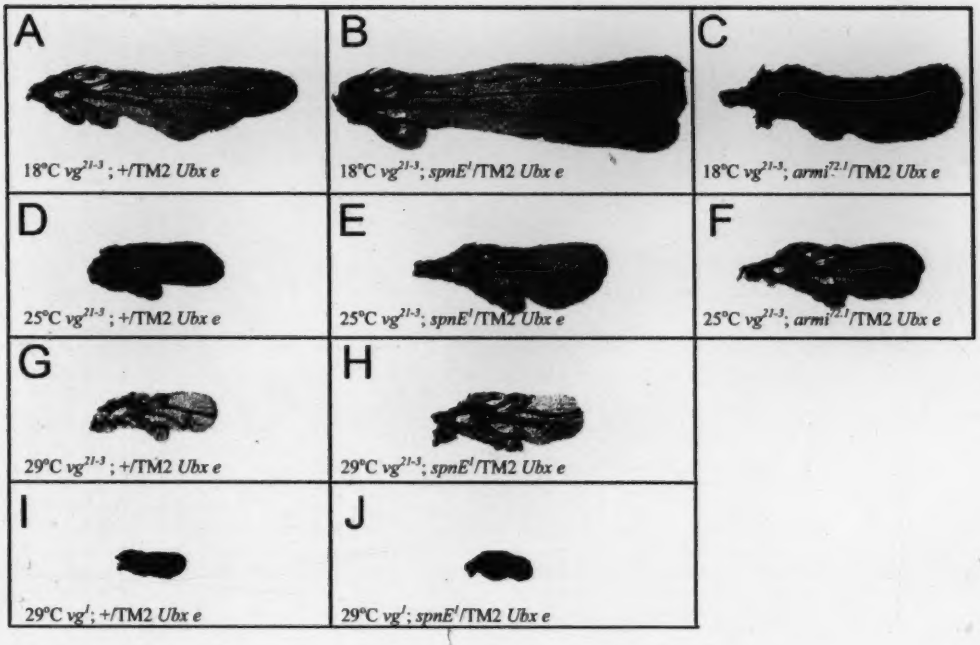

A
18°C vg^{21-3} ; +/TM2 *Ubx e*
B
18°C vg^{21-3}; $spnE^{1}$/TM2 *Ubx e*
C
18°C vg^{21-3}; $armi^{72.1}$/TM2 *Ubx e*
D
25°C vg^{21-3} ; +/TM2 *Ubx e*
E
25°C vg^{21-3}; $spnE^{1}$/TM2 *Ubx e*
F
25°C vg^{21-3}; $armi^{72.1}$/TM2 *Ubx e*
G
29°C vg^{21-3} ; +/TM2 *Ubx e*
H
29°C vg^{21-3}; $spnE^{1}$/TM2 *Ubx e*
I
29°C vg^{1}; +/TM2 *Ubx e*
J
29°C vg^{1}; $spnE^{1}$/TM2 *Ubx e*

Figure V-6: Wing phenotype of *vg* alleles over deletions at 25°C. Flies of each genotype consistently produced wings of the phenotypes shown. **A.** Homozygous $vg^{21\text{-}3}$ flies are strong mutants. **B.** Homozygous vg^{2a33} flies are mild mutants with some nicking of the wing margin. **C.** $vg^{21\text{-}3}$/ Df(2R)vg^{B} wings are slightly more severe than the wings of $vg^{21\text{-}3}$ homozygotes. **D.** vg^{2a33}/ Df(2R) vg^{B} are much more severe than their homozygous counterparts in **B**. **E.** $vg^{21\text{-}3}$/ Df(2R)CX1 wings are restored to a mild-to-moderate wing phenotype. **F.** vg^{2a33}/ Df(2R)CX1 wings are the same as with the Df(2R)vg^{B} deletion.

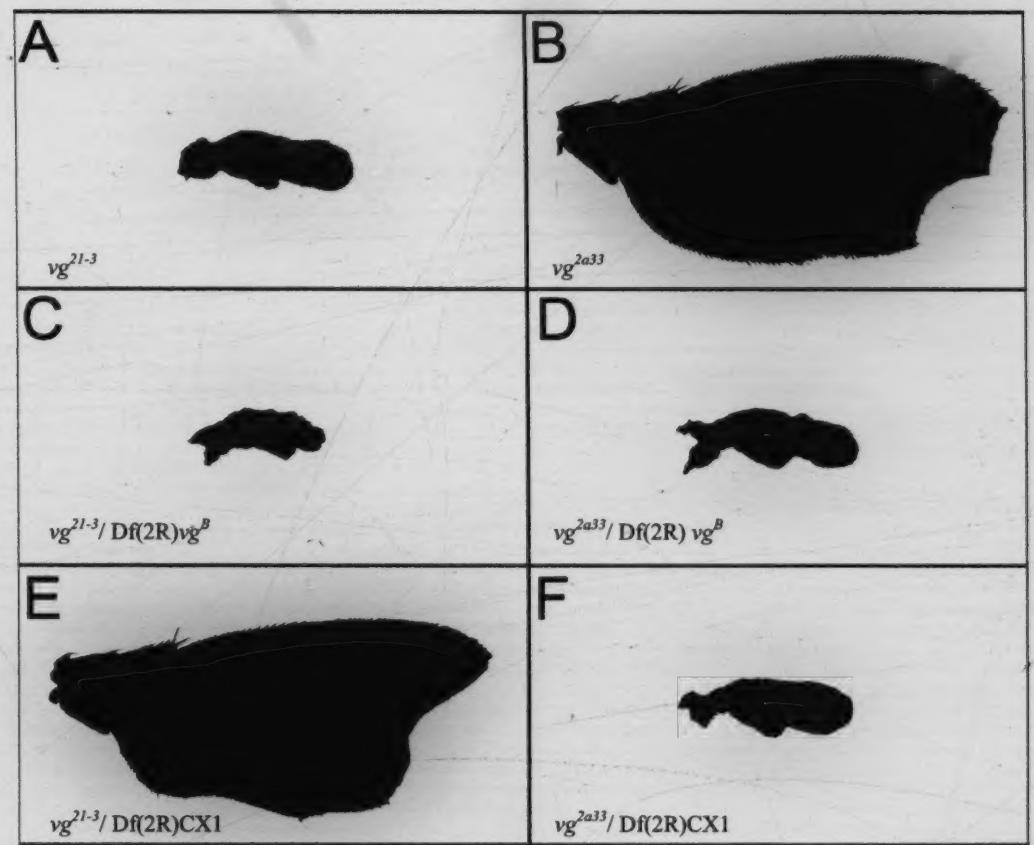
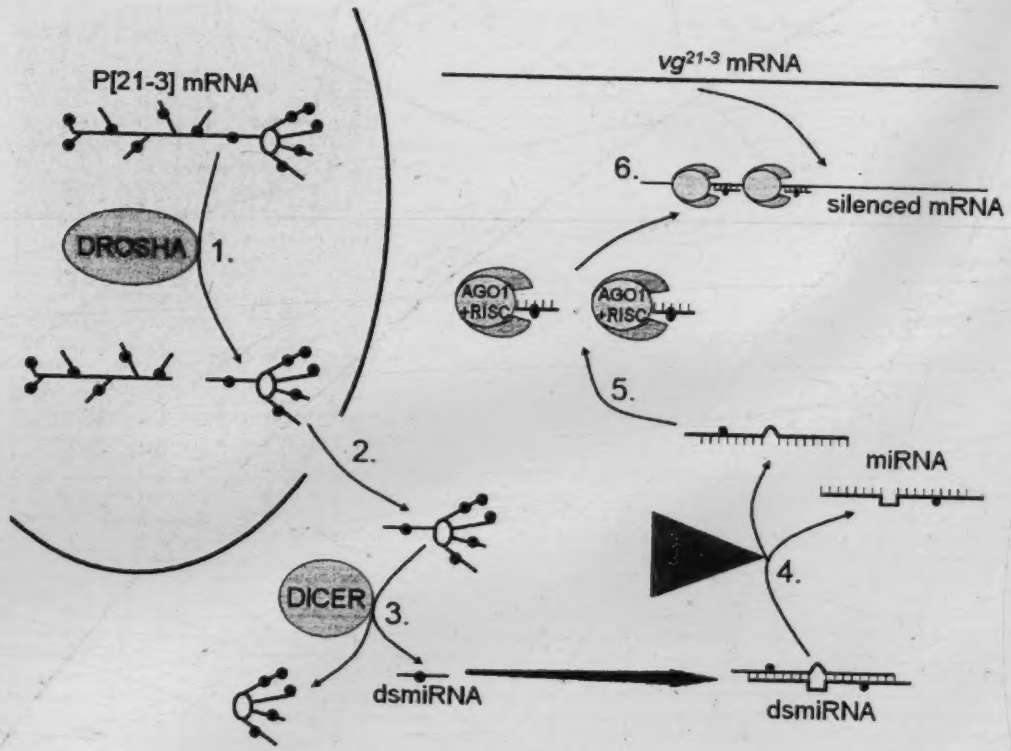

A
vg^{21-3}
B
vg^{2a33}
C
vg^{21-3}/ Df(2R)vg^{B}
D
vg^{2a33}/ Df(2R) vg^{B}
E
vg^{21-3}/ Df(2R)CX1
F
vg^{2a33}/ Df(2R)CX1

Figure V-7: Model of P[21-3] miRNA production and action. Step 1. The P[21-3] mRNA assumes a secondary structure in the nucleus which is recognized by the miRNA initiating RNAse Drosha to cleave one end of the mRNA away from the miRNA (HAMMOND *et al.* 2001). **Step 2.** The cleaved mRNA leaves the nucleus. **Step 3.** The RNAse Dicer cleaves the remainder of the mRNA away from the double stranded miRNA (LANDTHALER *et al.* 2004). The expanding arrow in step 3 is used to show the structure of the dsmiRNA in greater detail. **Step 4.** The double stranded miRNA is unwound by RNA helicases such as Armitage and Spindle E (PAL-BHADRA *et al.* 2004; TOMARI *et al.* 2004). **Step 5.** The miRNAs are now able to be loaded into the mature RISC through Argonaute 1(LEE *et al.* 2004). **Step 6.** The $vg^{21\text{-}3}$ mRNA cannot be translated as it is bound and silenced by the RISC:miRNA complex. In the presence of *P* element repressor, the production of P[21-3] mRNA would be halted and this would prevent accumulation of a silencing miRNA.


P[21-3] mRNA
DROSHA
1.
2.
DICER
3.
dsmiRNA
vg^{21-3} mRNA
6.
silenced mRNA
AGO1 +RISC
AGO1 +RISC
5.
miRNA
4.
dsmiRNA

Literature Cited

ANDERSON, K. J., M. M. DAVIS and R. B. HODGETTS, 2006 Novel events associated with phenotypic reversion of a P element mutant in Drosophila melanogaster. Genome **49:** 1184-1192.

BARTEL, D. P., 2004 MicroRNAs: genomics, biogenesis, mechanism, and function. Cell **116:** 281-297.

DANIELS, S. B., K. R. PETERSON, L. D. STRAUSBAUGH, M. G. KIDWELL and A. CHOVNICK, 1990 Evidence for horizontal transmission of the P transposable element between Drosophila species. Genetics **124:** 339-355.

GEYER, P. K., K. L. RICHARDSON, V. G. CORCES and M. M. GREEN, 1988 Genetic instability in Drosophila melanogaster: P-element mutagenesis by gene conversion. Proc Natl Acad Sci U S A **85:** 6455-6459.

GLOOR, G. B., C. R. PRESTON, D. M. JOHNSON-SCHLITZ, N. A. NASSIF, R. W. PHILLIS *et al.*, 1993 Type I repressors of P element mobility. Genetics **135:** 81-95.

HAMILTON, A. J., and D. C. BAULCOMBE, 1999 A species of small antisense RNA in posttranscriptional gene silencing in plants. Science **286:** 950-952.

HAMMOND, S. M., S. BOETTCHER, A. A. CAUDY, R. KOBAYASHI and G. J. HANNON, 2001 Argonaute2, a Link Between Genetic and Biochemical Analyses of RNAi, pp. 1146-1150.

HODGETTS, R. B., and S. L. O'KEEFE, 2001 The mutant phenotype associated with P-element alleles of the vestigial locus in Drosophila melanogaster may be caused by a readthrough transcript initiated at the P-element promoter. Genetics **157:** 1665-1672.

KARESS, R. E., and G. M. RUBIN, 1984 Analysis of P transposable element functions in Drosophila. Cell **38:** 135-146.

KAUFMAN, P. D., and D. C. RIO, 1992 P element transposition in vitro proceeds by a cut-and-paste mechanism and uses GTP as a cofactor. Cell **69:** 27-39.

KENNERDELL, J. R., S. YAMAGUCHI and R. W. CARTHEW, 2002 RNAi is activated during Drosophila oocyte maturation in a manner dependent on aubergine and spindle-E. Genes Dev **16:** 1884-1889.

KIDWELL, M. G., 1981 Hybrid dysgenesis in Drosophila melanogaster: the genetics of cytotype determination in a neutral strain. Genetics **98:** 275-290.

KIDWELL, M. G., 1983 Evolution of hybrid dysgenesis determinants in Drosophila melanogaster. Proc Natl Acad Sci U S A **80:** 1655-1659.

KIDWELL, M. G., J. F. KIDWELL and J. A. SVED, 1977 Hybrid dysgenesis in Drosophila melanogaster: a syndrome of aberrant traits including mutation, sterility and male recombination. Genetics **86:** 813-833.

LANDTHALER, M., A. YALCIN and T. TUSCHL, 2004 The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. Curr Biol **14:** 2162-2167.

LASKI, F. A., D. C. RIO and G. M. RUBIN, 1986 Tissue specificity of Drosophila P element transposition is regulated at the level of mRNA splicing. Cell **44:** 7-19.

LEE, C. C., Y. M. MUL and D. C. RIO, 1996 The Drosophila P-element KP repressor protein dimerizes and interacts with multiple sites on P-element DNA. Mol Cell Biol **16:** 5616-5622.

LEE, Y. S., K. NAKAHARA, J. W. PHAM, K. KIM, Z. HE *et al.*, 2004 Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell **117:** 69-81.

LEMAITRE, B., and D. COEN, 1991 P regulatory products repress in vivo the P promoter activity in P-lacZ fusion genes. Proc Natl Acad Sci U S A **88:** 4419-4423.

LEMAITRE, B., S. RONSSERAY and D. COEN, 1993 Maternal repression of the P element promoter in the germline of Drosophila melanogaster: a model for the P cytotype. Genetics **135:** 149-160.

LIM, L. P., M. E. GLASNER, S. YEKTA, C. B. BURGE and D. P. BARTEL, 2003a Vertebrate microRNA genes. Science **299:** 1540.

LIM, L. P., N. C. LAU, E. G. WEINSTEIN, A. ABDELHAKIM, S. YEKTA *et al.*, 2003b The microRNAs of Caenorhabditis elegans. Genes Dev **17:** 991-1008.

NITASAKA, E., and T. YAMAZAKI, 1988 Excision of one of two defective P elements as the cause of alternate mutational events (sn+ and sne) of the singed bristle allele snw in Drosophila melanogaster. Jpn J Genet **63:** 303-312.

OKAMURA, K., A. ISHIZUKA, H. SIOMI and M. C. SIOMI, 2004 Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev **18:** 1655-1666.

PAL-BHADRA, M., B. A. LEIBOVITCH, S. G. GANDHI, M. RAO, U. BHADRA *et al.*, 2004 Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. Science **303:** 669-672.

ROBERTSON, H. M., and W. R. ENGELS, 1989 Modified P elements that mimic the P cytotype in Drosophila melanogaster. Genetics **123:** 815-824.

SCHWARZ, D. S., G. HUTVAGNER, B. HALEY and P. D. ZAMORE, 2002 Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. Mol Cell **10:** 537-548.

STAVELEY, B. E., T. R. HESLIP, R. B. HODGETTS and J. B. BELL, 1995 Protected P-element termini suggest a role for inverted-repeat-binding protein in transposase-induced gap repair in Drosophila melanogaster. Genetics **139:** 1321-1329.

STAVELEY, B. E., R. B. HODGETTS, S. L. O'KEEFE and J. B. BELL, 1994 Targeting of an enhancer trap to vestigial. Dev Biol **165:** 290-293.

TANG, G., 2005 siRNA and miRNA: an insight into RISCs. Trends Biochem Sci **30:** 106-114.

TOMARI, Y., T. DU, B. HALEY, D. S. SCHWARZ, R. BENNETT *et al.*, 2004 RISC assembly defects in the Drosophila RNAi mutant armitage. Cell **116:** 831-841.

WILLIAMS, J. A., and J. B. BELL, 1988 Molecular organization of the vestigial region in Drosophila melanogaster. Embo J **7:** 1355-1363.

WILLIAMS, J. A., S. S. PAPPU and J. B. BELL, 1988a Molecular analysis of hybrid dysgenesis-induced derivatives of a P-element allele at the vg locus. Mol Cell Biol **8:** 1489-1497.

WILLIAMS, J. A., S. S. PAPPU and J. B. BELL, 1988b Suppressible P-element alleles of the vestigial locus in Drosophila melanogaster. MGG **212:** 370-374.

ZAMBON, R. A., V. N. VAKHARIA and L. P. WU, 2006 RNAi is an antiviral immune response against a dsRNA virus in Drosophila melanogaster. Cell Microbiol **8:** 880-889.

ZUKER, M., D. MATHEWS and D. TURNER, 1999 Algorithms and thermodynamics for RNA secondary structure prediction in *A practical guide in RNA biochemistry and biotechnology*, edited by J. BARCISZEWSKI and B. CLARK. Kluwer Academic Publishers, Dordrecht, Holland.

Chapter VI: Conclusions.

The research presented in this thesis adds substantially to our current understanding of the repressor encoding potential of *P* elements (Chapters III and IV), and reveals a novel mechanism of gene impairment using *P* encoded miRNA (Chapter V). Taken together, these two lines of research offer a better understanding of the apparent ease with which *P* elements became ubiquitous in natural *Drosophila melanogaster* populations worldwide during the 20th century. The wide range of phenotypes and different responses to repressor observed from *P* element alleles of *vestigial* can only be explained by the sequence within each *P* element. To make sense of how each element modulates the phenotype, we created additional mutants, the analysis of which allowed us to understand some of the sequence requirements for phenotypic severity and *P* repressor suppression.

Novel repressor-producing elements

The ability of the *P* element to limit its own transposition by the production and accumulation of repressor protein is a marvel of natural selection. If the *P* element were unable to repress itself, mutations would continue to accumulate in the host species until the collective phenotypic load rendered the individuals sterile or non-viable. The situation on the opposite end of the spectrum would be a transposon which was so successful at limiting transposition that it could not gain a sufficient presence in a species' genome to spread successfully. The *P* element was able to go from undetectable levels in wild populations, to being ubiquitously present in less than 50 years (ANXOLABEHERE *et al.* 1988; DANIELS *et al.* 1990; KIDWELL 1983). This rapid, but not

catastrophic, genome invasion would suggest that *P* elements are very effective at producing repressor protein once they are established in a population. Transposition frequently creates *P* elements with internal deletions created by aborted repair of the break left at the transposition site (GLOOR *et al.* 1991; O'HARE and RUBIN 1983). The production of internally deleted elements can accelerate the repression of transposition as the *P* elements are unable to make transposase, but may retain repressor producing sequences. Previous studies on the repression capacity of incomplete *P* elements have led to the development of two categories of repressor encoding elements, Type I and Type II elements (GLOOR *et al.* 1993; RASMUSSON *et al.* 1993). The Type I element category was experimentally specified as any *P* element which has a minimum of bases 1 – 1956 intact. Successive deletions 5' to base 1956 eliminated the ability to repress transposition. The Type II class of repressor producing elements is a catch-all group which has not been experimentally defined, but encompasses commonly found repressor elements found in natural populations. The KP element is the most prevalent of these mid-sized elements and contains a deletion of bases 808 - 2560 and produces a protein 37% of the length of a Type I repressor (BLACK *et al.* 1987; GLOOR *et al.* 1993). We have demonstrated that elements which do not fit into these groups can exhibit some properties of repressor-producing elements. The P[21-3] element falls less than 100 bases short of a Type I repressor, but was shown to repress *P* element expression in wing tissue using vg^{21-3} and in eye tissue using P[LacZ w^{+}] *ci* (Chapter IV). It is possible that the definitive breakpoint at 1956 for Type I repressors affects mRNA or protein stability, and slightly larger deletions like the one present in P[21-3] can restore repressor production. The P[21r36] element is also able to repress *P* element somatic expression at

just over half the size of a KP protein. This demonstrates that repressor activity does not rely on both the DNA binding THAP domain and the dimerization leucine zipper domain, but functional repression can be produced with just the THAP domain (Chapter IV). The discovery of these 2 elements which fall outside the published P repressor categories prompted us to propose that *P* element repressors are named according to the components which are present in their encoded proteins. The THAP, LZIP and PSI notations with superscripts for tissues exhibiting repression for the individual insertion proposed in Chapter IV are a more explicit and adaptable method for conveniently expressing to other researchers the nature of any *P* element under study.

microRNA co-suppression

The success of *P* elements in saturating the *Drosophila melanogaster* genome through recent horizontal transfer has occurred because of an efficient cut and paste mode of transposition, and an ability to minimize the creation of deleterious mutations. Insertions at *vestigial* reveal the types of mutations that can be caused by *P* elements and the effect of multiple *P* elements interacting on one phenotype. The molecular similarity of these alleles allowed us to focus on the subtle differences between them and draw conclusions about why each insertion produced the phenotype it did. Previous studies have suggested that the transcript coming from the *P* element promoter is the cause of mutant phenotypes in *vg* alleles (HODGETTS and O'KEEFE 2001; STAVELEY *et al.* 1995). Our results have shown that the $vg^{21\text{-}3}$ and vg^{2a33} alleles produce decreased levels of *P* element transcript in the presence of P repressor (Figure V-2), but while this reverts the $vg^{21\text{-}3}$ phenotype, it does not modify the vg^{2a33} phenotype (HODGETTS and O'KEEFE

2001). The fact that a reduction in *P*-initiated transcript in $vg^{21\text{-}3}$ will alleviate the phenotype is mirrored by the finding that RNAi mutants will also partially restore the wing phenotype (Figures V-4, V-5 and V-6). Taken together, these continue to support the hypothesis that *P* element transcription is detrimental in $vg^{21\text{-}3}$. The results of vg^{2a33}, however, do not support the *P* element transcriptional poisoning hypothesis as neither a reduction in *P* transcription, nor the presence of an RNAi mutant will modify the phenotype (Figure V-6).

The discovery that P[21-3] potentially creates a microRNA to downregulate similar mRNAs in the genome, such as the *vestigial* transcript with *P* element sequence in the 5' UTR, arose because of the availability of other *vg* alleles for comparison. This is the first evidence of the existence of such *P* element control of host gene function and it may have far-reaching consequences. The drive towards producing P cytotype in a newly invaded population is required to reduce transposition, but one consequence of the presence of repressor may be to reduce the genetic load on the host. If any of the new *P* element insertions act through the microRNA pathway to downregulate their associated genes, the transcriptional repression of those *P* elements by P repressor could restore gene function. Once *P* elements had firmly established themselves in the *Drosophila melanogaster* genome there would be selective pressure to maintain repressor-making elements as their loss would reactivate any latent miRNA production in addition to allowing for the creation of new mobilizations if a transposase-encoding *P* element were present.

The finding that the phenotype of $vg^{21\text{-}3}$ can be modulated by RNAi pathway mutants bears similarity to the silencing of retrotransposons caused by a newly described

RNA silencing pathway of Piwi-interacting, or piRNAs (LAU *et al.* 2006). It is important to note, however, that the piRNA pathway only operates in the germline (PELISSON *et al.* 2007; VAGIN *et al.* 2006 and reviewed in LIN 2007), and does not involve *Argonaute 1* or *Argonaute 2*, but instead the *Drosophila* piwi family members *Aubergine*, *Piwi*, and *Argonaute 3* (GUNAWARDANE *et al.* 2007). As the phenotype of $vg^{21\text{-}3}$ appears in somatic cells and is modified by loss of *Ago1* (Chapter V), we feel the involvement of the miRNA pathway in the production of a wing phenotype is much more likely than the piRNA pathway. It is possible, however, that the germline *P* element repression by P[lacZ] insertions at subtelomeric site 1A (RONSSERAY *et al.* 2003; RONSSERAY *et al.* 1993), which is known to be modified by *Aubergine* mutations (REISS *et al.* 2004) is created by piRNAs.

Further experiments

The phenotypic differences within the set of alleles, all inserted into the same position and in the same orientation at *vg*, cannot be explained by a single mechanism. The allele vg^{21} contains a large internal deletion, an intact promoter and set of 31 bp and 11 bp repeats. It is too short to express a protein with a complete THAP domain (Figure IV-8), but its strap-wing phenotype when heterozygous over a null allele reverts fully to wild type in the presence of repressor. In contrast, although vg^{21r36} contains very similar deletion endpoints (Figure III-3), it encodes a complete THAP domain (Figure IV-8). Like vg^{21}, it is homozygous wild type, but unlike vg^{21} when heterozygous over a null it gives a strong phenotype (Figure III-4). When comparing the secondary structure of transcripts produced by the two *P* elements by putting them through the mFOLD program

(Figure VI-1), it is noted that there are no extensive stem loops of the minimum 21 nucleotide size to encode a miRNA, and the smaller structures which are present are conserved between both transcripts. If the difference in phenotypic severity is due to the production of microRNA, then the structure producing that microRNA must be within the unique sequence for each element (highlighted regions in Figure VI-1). Prediction of RNA structures leading from the *P* elements into the *vestigial* region could be done to look for pairing between the unique P[21] and P[21r36] regions and *vestigial* exons, but the complexity of the sequence involved makes it difficult to produce a meaningful result. A more informative test for vg^{21r36} would be to observe it in the presence of a null allele and an RNAi mutant to assay the effect on wing phenotype. The most informative cross would use Df(2R)CX1 as the chromosome is deficient for both the *vestigial* and *Argonaute 1*, but *spindle E* or *armitage* mutants could also be used in combination with $Df(2R)vg^{B}$. It is unlikely that the phenotype of vg^{21} is mediated through a microRNA as there is such a mild loss of VG function, even when heterozygous with a null allele. With the prediction of miRNA producing regions comes the ability to directly assay for small RNA species containing those sequences in our mutant lines. Detection of microRNA species can be accomplished by a nuclease protection assay similar to the one used in this thesis to detect *P* element and *vg* transcripts. The basic protocol is to produce an internally labeled RNA probe which surrounds the suspect miRNA sequence, this probe is then hybridized in solution with total RNA and digested with an RNAse to remove all unprotected nucleotides from the probe. The sample would then be run on a denaturing acrylamide gel and if there is probe present which is sized between 18 and 24 nucleotides it would have been protected by the presence of a miRNA. It would be

expected that full length probe would also be found in this assay as it would be protected by the uncleaved transcripts which contain *P* element sequence.

Analysis of vg^{21}, $vg^{21\text{-}3}$ and vg^{2a33} phenotypes suggest that there are at least two levels of impairment created by *P* element *vg* alleles. A high degree of impairment for $vg^{21\text{-}3}$ is mediated extrinsically through the creation of microRNA which will downregulate both *P* and *vg* transcripts which contain complementary sequence to the microRNA. The low degree of impairment is produced intrinsically and can be observed when vg^{21} or $vg^{21\text{-}3}$ are placed in combination with a null allele and a strong P repressor such as P[Sal]89D (Figure III-1B and F). In both cases we see a very mild mutant wing phenotype which would suggest that there is a slight decrease in VG production. It is possible that the retention of *P* element sequence in the *vg* mRNA 5' untranslated region decreases the stability or translatability of the message enough to cause a mild wing phenotype. An extension of this idea could explain how an allele like vg^{2a33}, which does not respond to repressor, creates a mutant wing phenotype. The *P* element sequence in the vg^{2a33} transcript is missing the 3' 11 bp repeat, and this could cause greater RNA misfolding of the *vestigial* transcript and the resulting moderate wing phenotype.

While the production and study of new *P* element *vg* alleles is useful for creating hypotheses on how *P* elements can impair gene function, they do not easily allow for direct testing of those theories through production and testing of desired alleles. I believe that the systematic study of *P* element insertions on *vg* function should be moved into a system more suited to the testing of desired constructs. A system which utilized *Drosophila* cell culture would be much more amenable to the production of transgenic and modified constructs and a more rapid assay of gene function. A construct could be

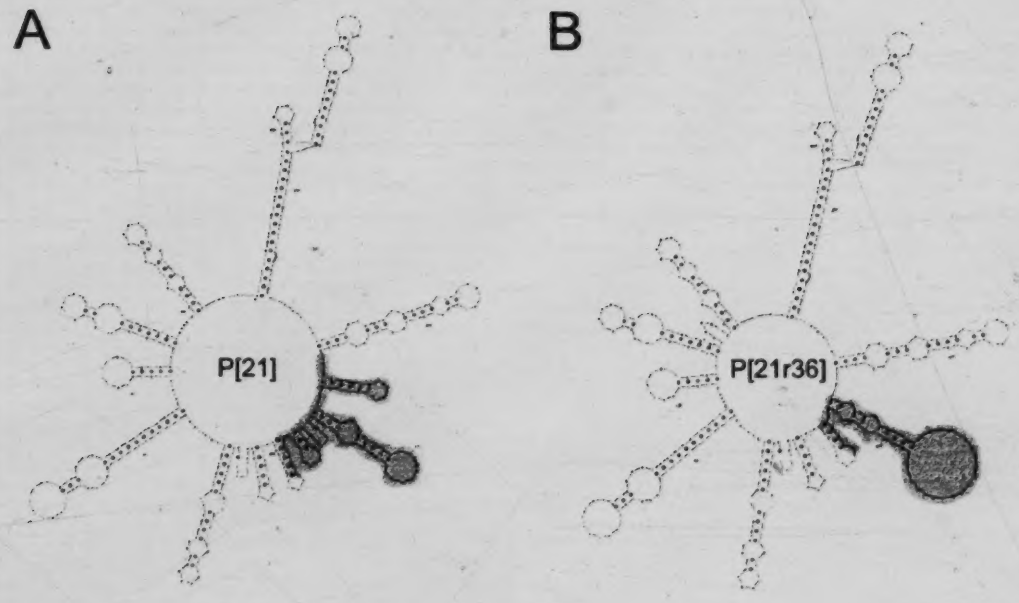
made in which the *vg* 5' UTR, containing a *P* element insertion, was placed upstream of an easily quantifiable reporter such as GFP. This construct would be driven by a promoter more highly expressed than the endogenous *vg* promoter to facilitate study in cell lines. Different versions of this construct could then be tested with known *P* element insertions, such as P[21] or P[21-3], to observe if loss of GFP expression mimics the loss of VG expression in our mutant stocks. If this system proved viable, it could then allow the testing of directed deletions of *P* element sequence without the need to select for rare replacement transposition events through microinjection. This system could not only be used for deletion studies, but more subtle alterations of individual bases. Targets such as those at the *P* element promoter, or in a putative microRNA encoding site, could be modified to assay their effect on gene expression.

Future implications

It is widely accepted that it is not solely the number of genes in a genome, but also the expression patterns of those genes which will create the complexity of an organism. There is a growing body of evidence to support the hypothesis that rather than being genomic parasites, transposable elements may be a mutagenic force which helps to create that complexity (reviewed in VOLFF 2006). The process of integration of a transposable element into the host genome is called domestication. In the case of *P* elements, it has been shown that the THAP domain has been incorporated into host genes of various *Drosophila* species through exon shuffling into existing genes and the creation of new genes (QUESNEVILLE *et al.* 2005). It is foreseeable then, that over time the *P* elements in *Drosophila melanogaster* may become more integrated into the genome. The

DNA-binding properties of *P* element repressor and its target, the *P* element termini, could be modified to become a transcription factor and enhancer pair with the simple addition of an activation domain to the repressor protein. The stock of $vg^{21\text{-}3}$; P[Sal]89D could be seen as a primitive domestication event as the protein product of P[Sal] is required to allow functional expression of VG. If expression of P[Sal]89D were to be lost in a portion of the wing disc, we would expect to see a corresponding change in wing morphology. Events like the contributions of additional *P* elements to the wing phenotype of *P* element *vestigial* alleles studied in this thesis can give insight into the process of transposable element domestication which could take millions of years to complete.

Figure VI-1: *P* element transcript folding from P[21] and P[21r36]. The mRNA encoded by the *P* elements are shown with unique DNA unique to each element highlighted in blue. **A.** P[21] 418 bp transcript. **B.** P[21r36] 390 bp transcript.


A
P[21]
B
P[21r36]

Literature Cited

ANXOLABEHERE, D., M. G. KIDWELL and G. PERIQUET, 1988 Molecular characteristics of diverse populations are consistent with the hypothesis of a recent invasion of Drosophila melanogaster by mobile P elements. Mol Biol Evol **5:** 252-269.

BLACK, D. M., M. S. JACKSON, M. G. KIDWELL and G. A. DOVER, 1987 KP elements repress P-induced hybrid dysgenesis in Drosophila melanogaster. Embo J **6:** 4125-4135.

DANIELS, S. B., K. R. PETERSON, L. D. STRAUSBAUGH, M. G. KIDWELL and A. CHOVNICK, 1990 Evidence for horizontal transmission of the P transposable element between Drosophila species. Genetics **124:** 339-355.

GLOOR, G. B., N. A. NASSIF, D. M. JOHNSON-SCHLITZ, C. R. PRESTON and W. R. ENGELS, 1991 Targeted gene replacement in Drosophila via P element-induced gap repair. Science **253:** 1110-1117.

GLOOR, G. B., C. R. PRESTON, D. M. JOHNSON-SCHLITZ, N. A. NASSIF, R. W. PHILLIS *et al.*, 1993 Type I repressors of P element mobility. Genetics **135:** 81-95.

GUNAWARDANE, L. S., K. SAITO, K. M. NISHIDA, K. MIYOSHI, Y. KAWAMURA *et al.*, 2007 A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. Science **315:** 1587-1590.

HODGETTS, R. B., and S. L. O'KEEFE, 2001 The mutant phenotype associated with P-element alleles of the vestigial locus in Drosophila melanogaster may be caused by a readthrough transcript initiated at the P-element promoter. Genetics **157:** 1665-1672.

KIDWELL, M. G., 1983 Evolution of hybrid dysgenesis determinants in Drosophila melanogaster. Proc Natl Acad Sci U S A **80:** 1655-1659.

LAU, N. C., A. G. SETO, J. KIM, S. KURAMOCHI-MIYAGAWA, T. NAKANO *et al.*, 2006 Characterization of the piRNA complex from rat testes. Science **313:** 363-367.

LIN, H., 2007 piRNAs in the germ line. Science **316:** 397.

O'HARE, K., and G. M. RUBIN, 1983 Structures of P transposable elements and their sites of insertion and excision in the Drosophila melanogaster genome. Cell **34:** 25-35.

PELISSON, A., E. SAROT, G. PAYEN-GROSCHENE and A. BUCHETON, 2007 A novel repeat-associated small interfering RNA-mediated silencing pathway downregulates complementary sense gypsy transcripts in somatic cells of the Drosophila ovary. J Virol **81:** 1951-1960.

QUESNEVILLE, H., D. NOUAUD and D. ANXOLABEHERE, 2005 Recurrent recruitment of the THAP DNA-binding domain and molecular domestication of the P-transposable element. Mol Biol Evol **22:** 741-746.

RASMUSSON, K. E., J. D. RAYMOND and M. J. SIMMONS, 1993 Repression of hybrid dysgenesis in Drosophila melanogaster by individual naturally occurring P elements. Genetics **133:** 605-622.

REISS, D., T. JOSSE, D. ANXOLABEHERE and S. RONSSERAY, 2004 aubergine mutations in Drosophila melanogaster impair P cytotype determination by telomeric P elements inserted in heterochromatin. Mol Genet Genomics **272:** 336-343.

RONSSERAY, S., T. JOSSE, A. BOIVIN and D. ANXOLABEHERE, 2003 Telomeric transgenes and trans-silencing in Drosophila. Genetica **117:** 327-335.

RONSSERAY, S., B. LEMAITRE and D. COEN, 1993 Maternal inheritance of P cytotype in Drosophila melanogaster: a "pre-P cytotype" is strictly extra-chromosomally transmitted. Mol Gen Genet **241:** 115-123.

STAVELEY, B. E., T. R. HESLIP, R. B. HODGETTS and J. B. BELL, 1995 Protected P-element termini suggest a role for inverted-repeat-binding protein in transposase-induced gap repair in Drosophila melanogaster. Genetics **139:** 1321-1329.

VAGIN, V. V., A. SIGOVA, C. LI, H. SEITZ, V. GVOZDEV *et al.*, 2006 A distinct small RNA pathway silences selfish genetic elements in the germline. Science **313:** 320-324.

VOLFF, J. N., 2006 Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. Bioessays **28:** 913-922.